Yangon University of Economics

Post Graduate Diploma in Research Studies

(9th Batch)

DRS-312 Forecasting Methods

Third Quarter

Forecasting Methods

Contents

16 Forecasting

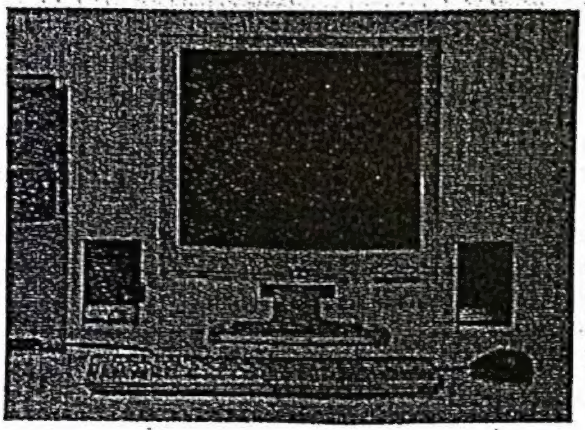
Learning Objectives

In this chapter, you learn:

- To appreciate the role of forecasting.
- To list and describe the qualitative and quantitative methods of forecasting
- To forecasting using time series models
- How to measure the forecast error to assess the accuracy of the models

G STATISTICS

@ Great Electronics



Imagine that you are the corporate planning manager of Great Electronics. One of your key tasks involves forecasting the sales of personal computers based on past internal as well as industry data. This exercise would involve a good understanding of time series data and the associated models for making reliable forecasts. For a couple of new products for which no past data are available, the forecasting exercise will have to be based on qualitative methods. The methods discussed in this chapter will help you achieve these aims in your corporate planning.

Many crucial decisions made by the management depend on the assessment o future demand for products and services, sales growth, and cost trends. The r agement must forecast to make sound decisions today. Hence, managers r efficient and reliable forecasting methods for business planning. This chapter presents a widely used forecasting techniques in practice.

16.1 Forecasting Basics

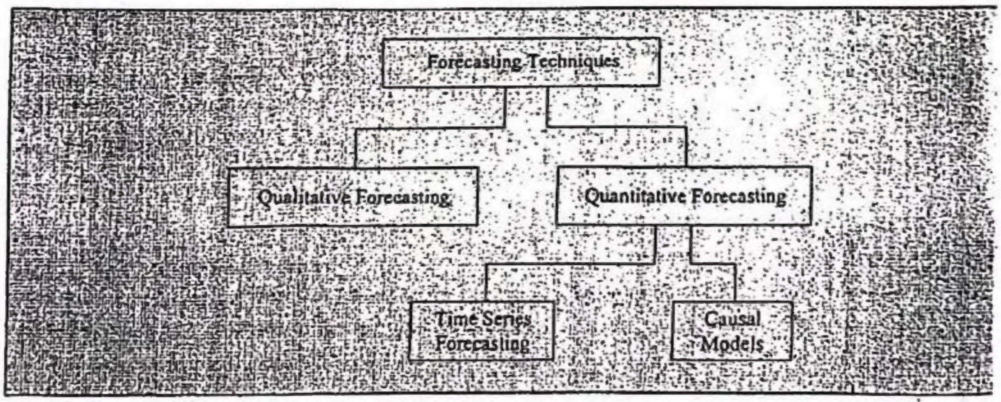
Why forecasting?

- Demand, or sales forecasting, is the foundation stone on which the entire business plan is built. An organization cannot predict its profitability without predicting sales reve Sales revenue cannot be predicted without forecasting sales in physical quantities. entire production program and materials resource planning cannot be achieved withc realistic sales forecast of the various products the organization would like to mai Corporate plans, turnaround plans, and competitive business strategies need the hel forecasting. In other words, not to forecast is to assume status quo and do nothing. This never be acceptable to any manager in any organization.
- We must, of course, recognize the fact that the future is uncertain and, therefore, no f casting can be hundred percent accurate. This is a paradox in forecasting; on one hand need sales forecasts and on the other hand no forecast can be accurate.
- Managers in any business enterprise have no choice between forecasting and no forecas because without a sound forecasting system, the risk of making a wrong decision incre Managers however have a choice amongst the methods of forecasting. In this chapter will dwell at length on the popular methods of forecasting in practice.

Forecasting Methods in Practice

Figure 16.1 portrays the widely used forecasting techniques in practice.

FIGURE 16.1
Forecasting Techniques in Practice



SELECTING THE RIGHT FORECASTING TECHNIQUE: GUIDELINES

- Availability of data: If no appropriate historical data are available, quantitative tec niques of forecasting are not possible. Only qualitative forecasting techniques are pos ble.
- Accuracy envisaged: Greater the accuracy needed, greater is the need for sophisticat techniques of forecasting.

- Urgency with which the forecast is sought. If forecasts are required urgently, only less sophisticated techniques are possible to use.
- Cost. This includes cost of forecasting exercise, and what it costs the firm if a wrong forecast is made.

.2 Qualitative Methods of Forecasting

Qualitative Forecasting

Assume that your company is about to introduce a new product that is unknown in the market. In this context, there will be no historical data that you could use to forecast your sales. It is a situation where you will find complete absence of any useful data. Under these circumstances, qualitative forecasting is the only method by which you could forecast sales for your new product.

Methods Used in Qualitative Forecasting

The following are the popular methods of qualitative forecasting. Four widely used methods of qualitative forecasting are

- Expert Opinion
- Market Survey
- Delphi Method
- Historical Analogy

PROGRESSIVE TEST QUESTION

Qualitative forecasting is superior to quantitative forecasting because it emphasizes on quality. State True or False.

SOLUTION The statement is false. These are the two broad types of forecasting methods. Qualitative is judgmental in nature whereas quantitative relies on hard data. By themselves these two have nothing to do with the word quality. Quality of the forecast is an entirely different subject and depends on many factors.

Expert Opinion: In this method, a group of experts from diverse backgrounds such as marketing, sales, finance, operations, and purchasing are asked to forecast sales of the product under consideration. A consensus is then reached on a forecast figure. Each expert brings with him or her a set of biases, and perspectives that might influence the forecast. Of course, their judgment would be substantiated by a wealth of information that includes past data, industry growth rates, competitive strategies and reactions from customers and distributors.

The advantages of this method are (1) it is fast and efficient; (2) it is timely and based on good information content; and (3) it uses the collective knowledge of experts.

The disadvantages of this method are (1) experts can make mistakes; (2) subjectivity and bias of experts can vitiate the forecast; and (3) the group dynamics of the experts could be greatly influenced by the degree of dominance of a particular person. He who shouts loudest, might get his way.

Market Survey: In this method, you conduct a market survey of a customer's intentions to buy a product. A carefully designed questionnaire is administered to the selected target audience of customers. Customers are selected independently using a representative random sample. This method is very popular and if carefully implemented will give you good results.

This is the most approprial technique to use, particularly if you want to forecast sales for a new product or new brand. This method of forecasting requires the active cooperation of the target audience. The sample size must be reasonably large. Larger the sample size, smaller will be the standard error and sampling error. Larger the sample size, the more time consuming and costly will the survey be. Hence, you have to strike a balance between sample size and cost.

Delphi Method: In the expert opinion method of forecasting, a consensus forecast is at after eliciting the opinion and views of experts with diverse backgrounds. Certainly method is subject to group dynamics (effects). At times, judgments may be highly influ persuasions of some group members who have strong likes and dislikes. The Delphi attempts to retain the wisdom and accumulated knowledge of a group while simultan attempting to reduce the group effects.

In the Delphi method, group members are asked to make individual assessment about a cast. These assessments are compiled and then fed back to the members, so that they opportunity to compare their judgment with others. They are then given an option to revise forecasts. After three or four replications, group members reach their final conclusion.

Historical Analogy: This method is applied when a new product is about to be introduc a company. Forecasting sales for new products is difficult in view of lack of proper histo data. Historical analogy method attempts to forecast sales for a new product based on the formance of related or similar products in the market place. The database of sales of these p ucts forms the basis for forecasting.

The drawbacks of this method include

You cannot predict, how your new product is similar or related to a particular prod Suppose you have a number of products that you feel are similar to yours. Which of these you consider as most similar to yours?

Products that are similar to yours could have failed in the past for a variety of reasons. us assume a similar product failed in the past because whenever there was an advertisen about this product, it was not available on the shelf. Hence, the consumers developed a nega perception about this product and became skeptical about its availability. You may not be av of these reasons and may simply conclude that your product will also fail!

PROGRESSIVE TEST QUESTION

Which one of the following will not affect the selection of a forecasting method?

1. Accuracy envisaged
2. Professional style of the organization
3. Data availability
4. Cost

SOLUTION The right choice is 2. Selection of a forecasting method will certainly depen 1, 3, and 4. Selection of a forecasting method will not depend on the professional style of organization.

PROGRESSIVE TEST QUESTION

What are the four methods of qualitative forecasting?

SOLUTION

1. Expert Opinion
2. Market Survey
3. Delphi Method
4. Historical Analogy

PROGRESSIVE TEST QUESTION

What is the paradox in forecasting?

SOLUTION Because the future is uncertain, no forecasting can be hundred percent accu This is a paradox in forecasting, on one hand you need sales forecasts and on the other han forecast can be accurate.

"If at all forecasting techniques would be useful, they would be only in predicting sales for products. They play an insignificant role in sales of services."

5.3 Quantitative Methods of Forecasting

Quantitative forecasting uses statistical analysis of data to forecast sales. Time series analysis and causal model fall under the purview of quantitative forecasting. In this chapter, we will discuss time series analysis (projective methods) of forecasting and causal model that uses regression Chapter 12 and 13 for forecasting. You are expected to revise Chapter 10, where we have already covered regression analysis extensively, so that you will be able to appreciate regression method of forecasting when we discuss it.

Time Series Analysis

Time series are series of observations that are taken at regular intervals of time. Data on weekly sales, monthly sales, and annual sales are examples of time series data.

Like many other data sets, if you have a time series data set, the first step in analyzing it is to draw a graph, particularly a simple scatter diagram or a line graph that will reveal sharply any underlying patterns.

Time series is made up of four components

- *Trend (T)* represents the long-term behavior of a time series. This would tell whether the time series data reveal a steady upward or downward movement.
- *Seasonal Variation (S)* represents variation caused by season. Typically, this shows variation in demand during peak and lean season. For example, demand for snow tires will be at its peak during winters in the United States.
- *Cyclical Variation (C)* represents the typical business cycles that occur sporadically in several years. For example, in the stock market, you will witness a cycle of buoyancy or boom and a cycle of recession that occurs once in a while between many years.
- *Random Variation (R)* represents irregular variations that occur by chance having no assignable cause. Random variation cannot be predicted.

Moving Average: The pattern revealed in observations varies over a time horizon. Instead of taking the average of all historical data, only the latest *n* periods of the data are used to get a forecast for the next period. This is the very essence of moving average forecast.

MOVING AVERAGE FORMULA

Moving Average (MA) Forecast for the next period

= Average of *n* most recent time series data (16.1)

XAMPLE 16.1
oving average

A company is interested in forecasting demand for one of its products. Past data on demand for the last 12 months are available and given below: Using a period of 3 months, forecast moving average forecast for period 13 (13th month).

Month	Sales (100 units)
1	15
2	9
3	16
4	17

5	11
6	20
7	10
8	17
9	12
10	9
11	18
12	20

SOLUTION

TABLE 16.1
Spreadsheet Showing the Moving Average Calculation for the Example Problem

Month	Sales (100 units)	3-Monthly Moving Average
1	15	—
2	9	
3	16	13.33
4	17	14.00
5	11	14.67
6	20	16.00
7	10	13.67
8	17	15.67
9	12	13.00
10	9	12.67
11	18	13.00
12	20	15.67

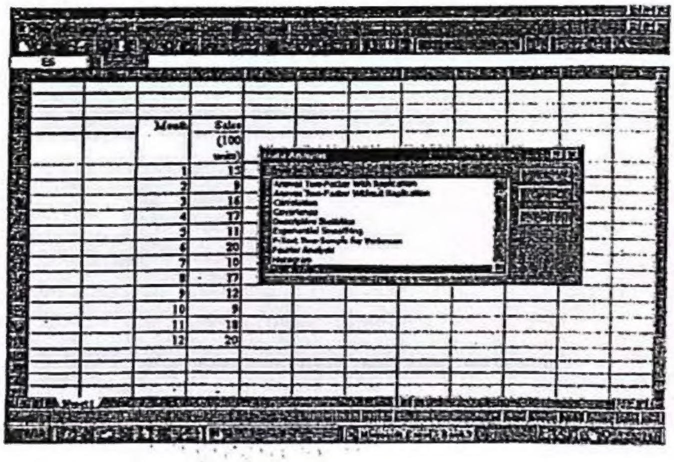
The first moving average is (15 + 9 + 16)/3 = 13.33. The second moving averag (9 + 16 + 17)/3 = 14.00, likewise, all entries are filled in this spreadsheet. As you will notice, number of moving averages in column 3 are only 10 compared to the original number of ob vations of 12; hence by using moving average method, you have lost 2 observations. Th inevitable in the moving average method of forecast.

Moving Average Using Microsoft Excel

We will calculate the moving averages for Example 16.1 just discussed, using Microsoft E in Figure 16.2.

Step 1: Click tools, click Data Analysis, and then click Moving Average. You get:

FIGURE 16.2
Microsoft Excel Worksheet for Computing Moving Average



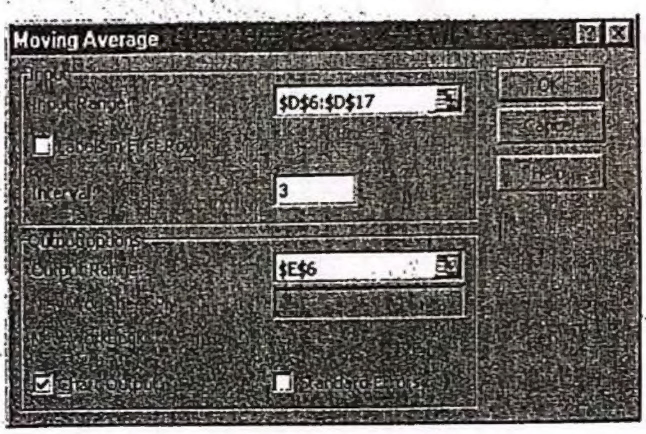
Step 2: Click OK, and then enter ***Input Range*** using the mouse by highlighting cells in column D from D6 to D17 in Figure 16.3. Enter 3 in the cell called ***Interval*** to indicate that 3 period (month) moving average is solicited. Specify the ***Output Range*** as E6. Click the ***Chart Output*** option to get the graph also. The screen shots are given in Figs 16.4 & 16.5.

Please note that if you do not specify ***Output Range***, Excel will give the output in a New Worksheet. So, you face no hassles. The flexibility and versatility provided by Excel is truly outstanding.

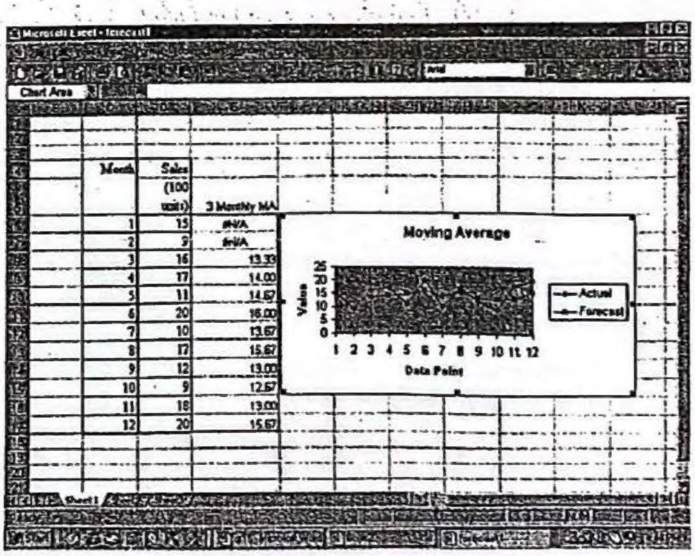
Step 3: Click OK and you get:

In column E, 3 month moving averages are worked out in Figure 16.4. Moving averages for months 1 and 2 are obviously not possible. The first moving average is (15 + 9 + 16)/3 =13.33. The second moving average is (9 + 16 + 17)/3 =14.00. Likewise all entries are filled in by Excel, automatically. As you will notice that the number of moving averages are only 10 in column E compared to the original number of observations of 12.

GURE 16.3
crosoft Excel
ving Average



GURE 16.4
crosoft Excel
rksheet for
mputing Moving
erage



The forecast demand for month 13 = 15.67 (moving average corresponding to month 12). Note that the forecast for the next period is always the most recent moving average.

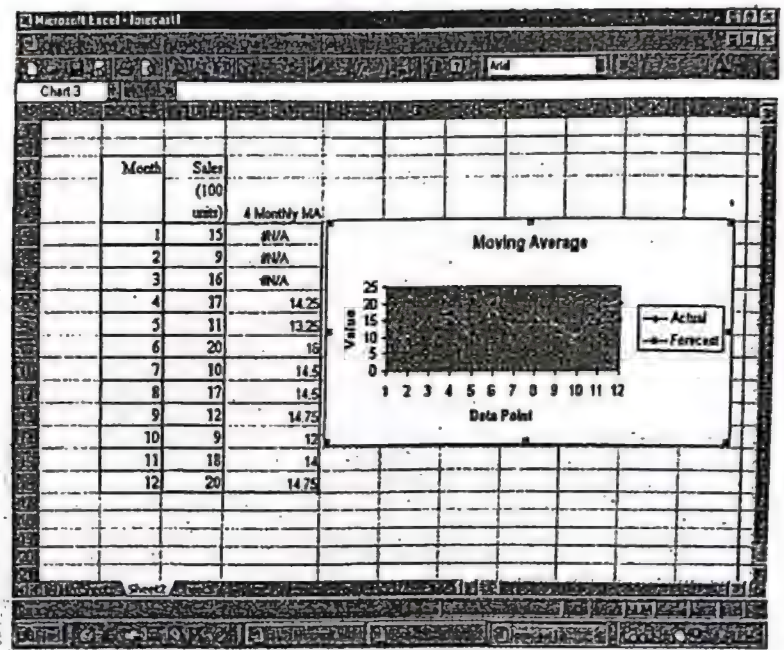
PROGRESSIVE TEST QUESTION

What is the moving average forecast demand for month 13 if you use a 4-monthly interval (period)?

SOLUTION Using Microsoft Excel in Figure 16.5, we have the following solution:

Forecast for month 13 is = 14.75 (Moving average corresponding to period 12 is 14.75).

FIGURE 16.5
Microsoft Excel Worksheet for Computing Moving Average Forecast



Month	Sales (100 units)	4 Monthly MA
1	15	#N/A
2	9	#N/A
3	16	#N/A
4	17	14.25
5	11	13.25
6	20	16
7	10	14.5
8	17	14.5
9	12	14.75
10	9	12
11	18	14
12	20	14.75

Drawbacks of Moving Average Forecast

Moving average forecast is quick, easy, and fairly inexpensive to implement. It provides a rea sonably good forecast for the immediate future (very short term). However, practicing manager must remember the drawbacks

- Moving averages do not react well to seasonal variations
- All observations considered in a time horizon are given the same weight
- A large amount of historical data should be gathered and maintained to update forecast value
- The choice of the period (n) is generally arbitrary.

Exponential Smoothing Exponential smoothing is a particular case of moving average in whic there are three components: (1) the forecast for the most recent period; (2) the actual value fo the period; and (3) a smoothing constant α. This smoothing constant α is a weighting factor tha lies between 0 and 1. The selection of the right kind of α is matter of judgment by the experi enced user. It must be chosen carefully.

Exponential smoothing is an excellent forecasting technique for short-term forecasting. It i used not only in sales forecasting but also in forecasting input prices in materials procuremen The single biggest advantage is that this technique is extremely simple to use.

EXPONENTIAL SMOOTHING FORMULA

$$\text{New Forecast} = \alpha(\text{actual value}) + (1 - \alpha)(\text{last forecast}) \quad (16.2)$$

The meaning of this statement is explained with an example. Suppose, the actual sale for month 2 is 50 units and your forecast for month 2 is 55 units. Let us take $\alpha = 0.3$.

$$\text{New Forecast} = (0.3)(50) + (1-0.3)(55) = 53.5$$

Exponential smoothing is a versatile technique of forecasting that allows the user a great deal of flexibility. You can choose the right α over a period of time using your experience. You can decide how much weight you should give to the recent actual value, and how much to the forecast based on your experience of the recent past.

XAMPLE 16.2

xponential moothing

A company is interested in forecasting the demand for one of its products. Past data on demand for the last 12 months are available and given below: Using exponential smoothing technique, forecast demand for month 13. Take $\alpha = 0.2$.

Month	Sales (100 units)
1	15
2	14
3	16
4	17
5	15
6	18
7	20
8	22
9	23
10	21
11	24
12	26

SOLUTION

ABLE 16.3

preadsheet Showing asic Calculations

		$\alpha = 0.2$
Months	Sales (100 units)	Smoothing Values
1	15	
2	14	15.00
3	16	14.80
4	17	15.04
5	15	15.43
6	18	15.35
7	20	15.88
8	22	16.70
9	23	17.76
10	21	18.81
11	24	19.25
12	26	20.20

The smoothing values in the 3rd column are calculated using the formula, New Foreca α(actual value) + $(1-\alpha)$(last forecast). Just for clarity, let me show a couple of smoothed ues. There will be no smoothing value possible for the first month. For the second month, smoothed value is taken as the previous month's actual. In this case, it is 15. Now, applying formula,

$$\text{New Forecast} = 0.2(14) + (1-0.2)(15) = 14.8$$

This is the forecast that appears in column 3 against month 3. For the 4th month, again ap the formula,

$$\text{New Forecast} = 0.2(16) + (1-0.2)(14.8) = 15.04$$

This is the forecast for the 4th month. Proceeding in this manner, all the exponen smoothing forecasts for the remaining months can be worked out.

$$\begin{aligned}\text{Forecast for the 13th Month} &= 0.2(\text{actual value}) + (1-\alpha)(\text{last forecast})\\ &= 0.2(26) + (1-0.2)(20.2) = 21.36\end{aligned}$$

This is the demand forecast for the month 13.

EXAMPLE 16.3

Exponential Smoothing Using Microsoft Excel

Step 1: In Microsoft Excel spreadsheet (Figure 16.6), click Tools, click *Data Analysis*, and cl *Exponential Smoothing*. The following screen will appear.

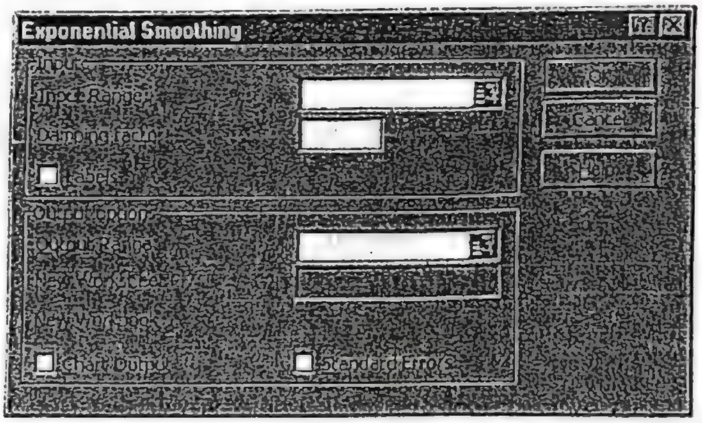
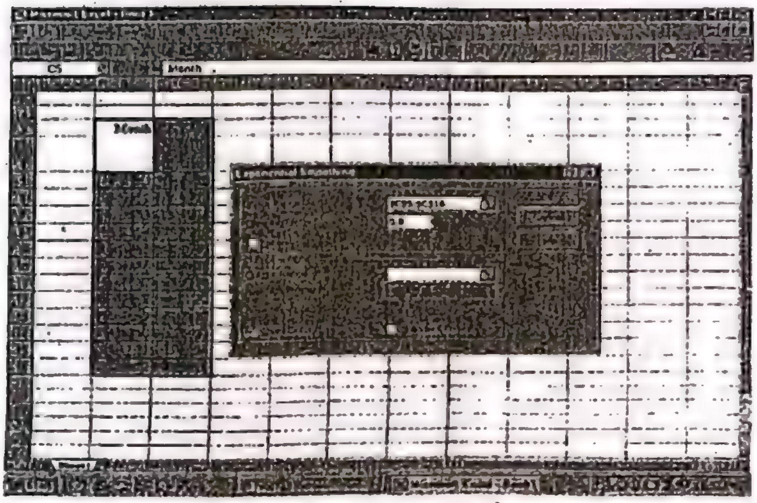


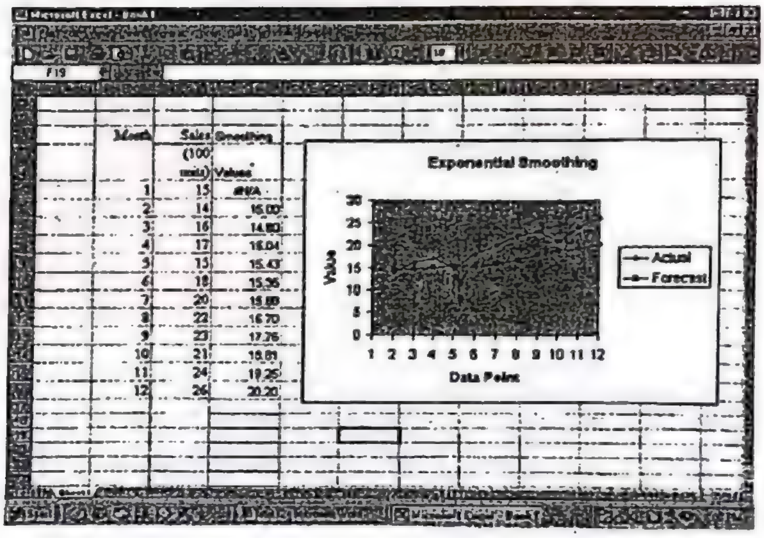
FIGURE 16.6
Excel Screen shots for Data Entry.

Step 2: In the prompt, highlight as usual Input Range. In the cell for *Damping factor*, ent 0.8 (This is Excel's way of asking for the value of α). Note that you have been giv $\alpha = 0.2$. Highlight range for output, and then click OK in Figure 16.7. Before clicking OI you can also click for *Chart Output* so that you obtain the graph of forecast values wi actuals.



GURE 16.7
crosoft Excel
rksheet for
mputing
onential Smoothing

Step 3: Click OK, for results in Figure 16.8.



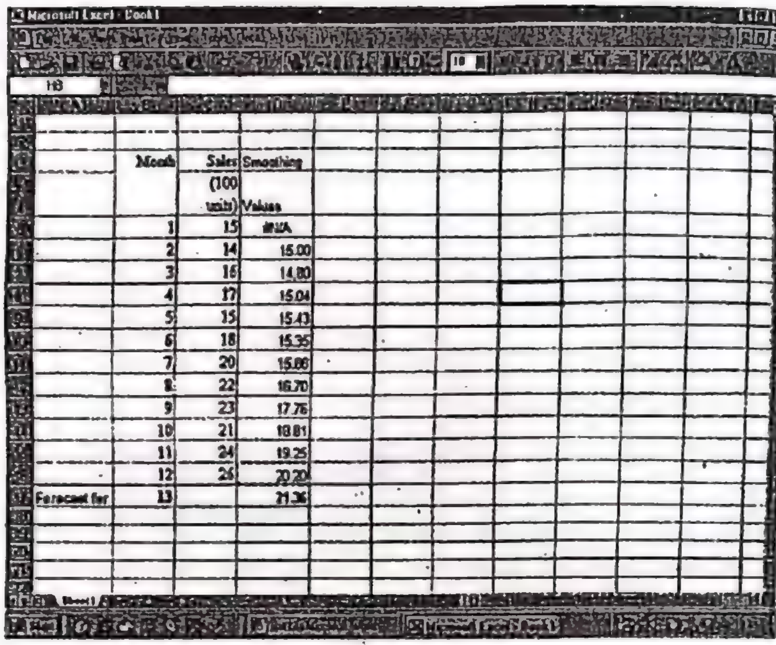
URE 16.8
el out for
onential
oothing

This is exactly identical to the results you have obtained by actual calculation method. See exponential smoothing forecast values that appear in column D.

To forecast for the month of 13, bring the mouse to cell D16 and click. Then click *Edit*, click *Copy*. Bring the mouse to cell D17 and then click *Paste* icon. The answer will appear in cell D17. This answer screen is shown in Figure 16.9.

Scanned with CamScanner

FIGURE 16.9
Excel Output-Exponential Smoothing



	Month	Sales (100 units)	Smoothing Values
	1	15	#N/A
	2	14	15.00
	3	16	14.80
	4	17	15.04
	5	15	15.43
	6	18	15.35
	7	20	15.88
	8	22	16.70
	9	23	17.76
	10	21	18.81
	11	24	19.25
	12	26	20.20
Forecast for	13		21.36

You can see in this Excel output, forecast for the month of 13, appears in cell D17 in b The answer shown is 21.36. This is the same as what you had obtained by the actual calcula in Example 16.2 using the Equation (16.2).

POINTS TO PONDER ON SELECTION OF SMOOTHING CONSTANT

The question that is often raised in the usage of exponential smoothing is: "Is there a scienti way to fix the value of the smoothing constant α?" The answer is both, *Yes* and *No*. Y because you can get that the value of α that gives the minimum mean square err No, because mean square error is highly influenced by the square terms of individual erro Judgment of α based on experience and tracking forecasting efficiency is mandatory. This indeed a weakness of the exponential smoothing method of forecasting.

Trend Projection In this method, we fit a trend line using the time series data. This trend could be linear or nonlinear (quadratic trend, exponential trend, etc). We will discuss the li trend that is popular, and much used in practice. The reason for its popularity emerges from following rationale.

Most of the nonlinear trend lines can be converted into linear lines by mathematical tr formation. The linear trend is a reasonably good approximation of trend pattern that is reve by time series data.

In simple terms, we fit an equation of the form $Y = \alpha + bt$ using the method of least squ The least square method used in regression analysis that was discussed in Chapter 12. In equation, Y is the dependent variable and t is the independent variable. In other wo you assume that forecast values of Y will be shaped by the past pattern only. The historical tr continues.

:AMPLE 16.4 A company is interested in forecasting demand for one of its products. Past data on demand for the last 12 months are available and is given here Using linear trend line, forecast demand for month 13.

Month	Sales (100 units)
1	15
2	14
3	16
4	17
5	15
6	18
7	20
8	22
9	23
10	21
11	24
12	26

SOLUTION You have two options. Option 1 is to use the formula approach involving a simple linear regression model discussed in Chapter 12; the option 2 is to use Data Analysis of Excel.

The formula used in the least square approach are given below.

FORMULA FOR TREND PROJECTION OF THE FORM $y = a + bt$

$$b = \frac{\sum (t-\bar{t})(Y-\bar{Y})}{\sum (t-\bar{t})^2} \qquad (16.3)$$

$$a = \bar{Y} - b\bar{t}$$

The basic calculations are shown in Table 16.3.

BLE 16.3
sic calculation for
nd projection

Month	Sales (100 units)	$(t-\bar{t})$	$(Y-\bar{Y})$	$(t-\bar{t})(Y-\bar{Y})$	$(t-\bar{t})^2$
1	15	−5.5	−4.25	23.3750	30.2500
2	14	−4.5	−5.25	23.6250	20.2500
3	16	−3.5	−3.25	11.3750	12.2500
4	17	−2.5	−2.25	5.6250	6.2500
5	15	−1.5	−4.25	6.3750	2.2500
6	18	−0.5	−1.25	0.6250	0.2500
7	20	0.5	0.75	0.3750	0.2500
8	22	1.5	2.75	4.1250	2.2500
9	23	2.5	3.75	9.3750	6.2500
10	21	3.5	1.75	6.1250	12.2500
11	24	4.5	4.75	21.3750	20.2500
12	26	5.5	6.75	37.1250	30.2500
78	231			149.5000	143.0000

$\bar{t}$ and $\bar{Y}$ can be computed from the first two columns. $\bar{t} = 78/12 = 6.5$; $= 231/12 = 19.25$

The bottom row of the 5th and 6th columns represent

$$(t - \bar{t})\,(Y - \bar{Y}) \text{ and } (t - \bar{t})^2$$

Using Equations (16.3) and (16.4),

$$b = \frac{\sum (t-\bar{t})(Y-\bar{Y})}{\sum (t-\bar{t})^2} = \frac{149.50}{143.00} = 1.045$$

$$a = \bar{Y} - b\,\bar{t}$$

$$= 19.25 - 1.0455 * 6.5 = 12.45$$

Hence, the fitted line is given by,

$$Y = 12.45 + 1.045t$$

Forecast demand for month 13 = 12.45 + (1.045)(13) = 26.04 (units of 100)

To use the Data Analysis of Excel, Click Tools, click *Data Analysis*, and click *Regression*.

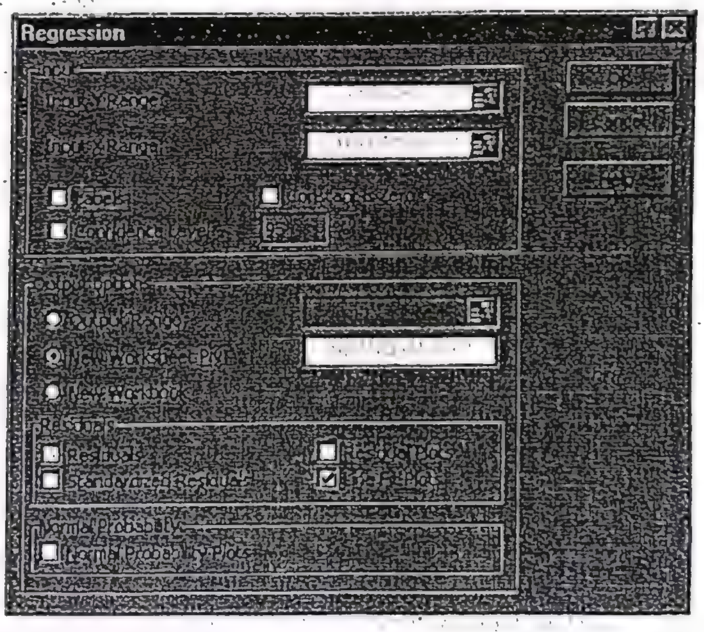
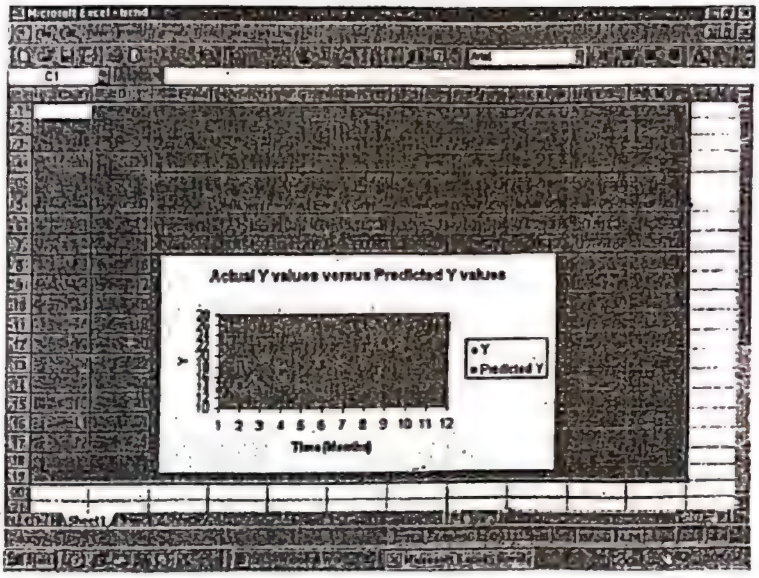


FIGURE 16.10
Excel initial Screen for Data Entry for Regression.

Enter data for Y range and X range just like what you have done in Chapter 12. Click with mouse *Line Fit Plots* and click OK, you get display as shown in Figure 16.11.

URE 16.11
el output fo the
ession problem



As you can see, intercept = 12.45 and slope = 1.045. Hence, the fitted line is given by $Y = 12.45 + 1.045t$ (Time t is the independent variable, but Excel has taken X as independent variable by default).

Forecast demand for month 13 = 12.45 + (1.045)(13) = 26.04

WHAT HAPPENS IF THE TREND EQUATION IS NOT LINEAR?

Two Nonlinear trend equations

1. Suppose the trend equation is of the form $Y = a + bt + ct^2$.
   How will you project future trend? Substituting $t^2 = u$. This will make the equation change to $Y = a + bt + cu$. This is a typical multiple regression model that can be solved using Microsoft Excel.
2. Suppose $Y = ae^{bt}$.
   Take log on both sides, this becomes $\log Y = \log a + bt$. This is of the form $Z = A + bt$, which is a simple linear regression model that can be solved using either formula approach or Microsoft Excel.

Forecasting Using Multiple Regression Model

Multiple Linear Regression Model Whenever you are interested in the combined influence of several independent variables upon one dependent variable that you want to forecast, your model is that of multiple regression. Demand, for example, may be a function of price, income of the consumer, advertising expense, industrial growth, and competitor's price. When all these independent variables change, what happens to the demand projection is a study of multiple linear regression.

STEPS TO BE CONSIDERED IN A MULTIPLE LINEAR REGRESSION MODEL

- Postulate the model $Y = a + bX1 + cX2 + dX3 + \ldots$
- Enter the sample data for Xs and Y in Microsoft Excel.
- Perform the Regression Analysis and get the summary output from Excel.

- Write the Regression Equation using the intercept and coefficient of Xs from Excel sum output. Predict Y for given Xs.
- Validate the model statistically by looking at R^2 as well as F statistic in the ANOVA tests the null hypothesis of no linear relationship.
- After statistical validation, use the model for estimation and prediction.

EXAMPLE 16.5
Sales Forcasting

To measure the effect of advertising and sales promotional efforts, the following data collected form a consumer marketing company for the last 10 months. Figures in the foll table are in $1000.

Month	Sales (Y)	Advertisement Expense ($X1$)	Sales Promotional Expense ($X2$)
1	200	45	15
2	250	50	20
3	300	55	24
4	650	85	45
5	400	65	30
6	300	55	25
7	320	57	27
8	450	68	32
9	350	60	28
10	550	70	37

Answer the questions, using Microsoft Excel:

1. Set up a regression model by taking Sales (Y) as the dependent variable and advertis expense ($X1$) and sales promotion expense ($X2$) as independent variables and valida model using R^2 value.
2. Forecast $X1$ and $X2$ for month 11 by using exponential smoothing technique.
3. Now forecast sales for month 11 using results obtained in 2.

SOLUTION 1 Invoke Data Analysis Pack of Microsoft Excel and enter the data as rec under Regression and execute the model. You get the following output:

(The details of each step are discussed in Chapter 12. So only the final output is given

Month	Sales (Y)	Advertisement Expense ($X1$)	Sales Promotional Expense ($X2$)
1	200	45	15
2	250	50	20
3	300	55	24
4	650	85	45
5	400	65	30
6	300	55	25
7	320	57	27
8	450	68	32
9	350	60	28
10	550	70	37

SUMMARY OUTPUT

Regression Statistics			*Coefficients*
Multiple R	0.987153		
R Square	0.974471	Intercept	195.761
Adjusted R Square	0.967177	X Variable 1	5.033
Standard Error	25.16275	X Variable 2	9.388
Observations	10		

Scanned with CamScanner

The output shows the following:

$$Y = 195.76 \text{ (2 places of decimal taken)}$$

$$X1 = 5.03,\ X2 = 9.39$$

1) The fitted model is

$$Y = 2195.76 + 5.03X1 + 9.39X2$$

The model has a good accuracy level as evident from the R^2 value that is quite high. $R^2 = 0.97$ (two places of decimal). Even the adjusted $R^2 = 0.97$, indicating the robustness of the model to predict. In other words, R^2 value is close to 1, and hence, the model is reliable in forecasting. This answers (part 1) of the question.

2) Invoke Exponential Smoothing under Data Analysis Pack. Enter the input data for $X1$ and $X2$ separately. Use a dampening factor of 0.7 (same as $\alpha = 0.3$). The following output emerges.

(As we have already covered the step-by step approach in this chapter, only the final output is given).

Exponential Smoothing Forecast Values

X1	X2
45.00	15.00
46.50	16.50
49.05	18.75
59.84	26.63
61.38	27.64
59.47	26.85
58.73	26.89
61.51	28.42
61.06	28.30

The smoothed values will start from period 2 only in Excel's exponential smoothing and will not be available for period 1.

Forecast of $X1$ for period 11 = α(actual value in period 10) + $(1 - \alpha)$(forecast of period 10)
= 0.3(70) + 0.7(61.06) = 63.74

Forecast of $X2$ for period 11 = α(actual value in period 10) + $(1 - \alpha)$(forecast of period 10)
= 0.3(37) + 0.7(28.3) = 30.91

3) Forecast of sales for period 11 is obtained by substituting the values of $X1$ and $X2$ (obtained in the previous part given above) in the regression equation

$$Y = -195.76 + 5.03X1 + 9.39X2$$

Forecast of sales for period 11 = 2195.76 + 5.03(63.74) + 9.39(30.91) = 415.1

Since sales are in $1000, the forecast for period 11 is a sale of $415100.

Limitations of Using Multiple Regression Model for Forecasting

The most crucial assumption made is that the independent variables are not correlated with each other. If they are correlated then the regression coefficients cannot be estimated. This problem is called **multicollinearity**. The procedure followed for resolving multicollinearity is to drop the independent variable that has the highest standard deviation and then rework the model again. You may also like to use two-stage least square method that is part of econometrics. The other way is to transform a set of correlated independent variables into an uncorrelated set of variables by the technique called **principal component analysis**. This is an advanced technique requiring the help of advanced statistical software like SPSS. When there are wild fluctuations in one or more of the independent variables, multiple regression model crumbles, and will be highly unreliable. To use the multiple regression model for prediction, you have to first predict the values of the independent

variables using some other prediction method. In forecasting problems, multiple regression, at l can work for short and medium term only. It cannot be successfully used for long-term forecasl

A Brief Note on Accuracy of Forecast

- We have discussed a number of forecasting techniques in this chapter. Needless to accuracy of forecast is paramount. Accuracy measures must reflect the closeness of dicted values with the actual values. Closer the predicted value to the actual, the greater accuracy. Backtracking ability of forecast is measured by the behavior of forecast va towards the actual. In all time series forecast methods that we have discussed so far, have provided graphical display of predicted versus actual values to understand the b tracking ability of forecast model under consideration.
- Another point to be noted is that suitable adjustments should be made in the forecast fig arrived at. This would include adjustments for seasonality and cycles. For example, have the trend projections based on least square line. You have made a projection for coming period. This projection figure will have to be suitably modified if there is a str seasonality. You can easily establish seasonal index for each calendar month. Thi obtained by dividing the actual value by the corresponding trend value. If you continuo maintain a large database, seasonal indices could be updated in a dynamic manner. All needs to be done is to first project the trend value for the forecast period by using the l square method, or moving average, and then multiply this trend value by the correspon seasonal index; you get a forecast adjusted for seasonality.
- All these measures will improve your accuracy. There are two methods in practice that used for understanding forecast error.

  1. Average Absolute Error This is obtained by computing the absolute difference betw forecast value and actual value for every element in the time series data set, and then tal the average of all these values.

  2. Average Percentage Relative Error In this method, you first compute the abso difference between forecast value and actual value for every element in the data set, then divide each one of them by the corresponding actual value. Take the average of s values and multiply by 100. You get average percentage error. The selection of one of th methods is a matter of judgment.
- Intuitively and logically, the graph of forecast values should be reasonably close to actual values. If it is not, look for reasons and gather more data. Revise your model c pletely, if needed.
- Accuracy can be greatly improved if you have a large amount of historical data. This permit you to use advanced forecasting techniques, like the Box-Jenkin method, Adap Filtering, and the Econometric models of forecasting. The discussion of these is beyond scope of this chapter. Those interested can refer to books on Econometrics for treatmen these advanced topics.

SUMMARY

This chapter has provided a conceptual framework on various forecasting techniques with their strengths and limitations. Specifically, this chapter is focused on:

- The need for forecasting.
- Schematic diagram giving classification of forecasting techniques in practice.
- Guidelines for selecting a forecasting method.
- Qualitative or judgmental forecasting split into expert opinion, market survey, Delphi method, and historical analogy.
- Quantitative forecasting split into time series anal and causal method. Detailed coverage on time se analysis as well as causal model involving regression
- Time series models split into moving average, exponer smoothing, and trend projection using least square line
- Forecasting accuracy and associated measures on f cast error.
- Use of Microsoft Excel to compute moving avera exponential smoothing, trend line based on least sq and multiple regression forecasting.

EY FORMULAS

ring Average Formula

ring Average (MA) Forecast for the next period
.verage of n most recent time series data. (16.1)

onential Smoothing Formula

/ Forecast = α(actual value)
+ (1 − α)(last forecast) (16.2)

Formula for Trend Projection of the form $Y = a + bt$

$$b = \frac{\sum (t - \bar{t})(Y - \bar{Y})}{\sum (t - \bar{t})^2} \tag{16.3}$$

$$a = \bar{Y} - b\,\bar{t} \tag{16.4}$$

EY TERMS

age absolute error 590
age percentage relative
rror 590
ical variation 577
ohi method 576
ert opinion method 575
onential smoothing 580
forecast error 590
historical analogy method 576
market survey method 575
mean square error 584
moving average 577
multicollinearity 589
principal component analysis 589
qualitative forecast 575
quantitative forecast 577
random variation 577
seasonal variation 577
time series 577
trend 577
trend projection 584

HAPTER REVIEW PROBLEMS

ecking Your Understanding

1 Moving average method of forecasting requires a
e amount of historical data. State True or False.

2 What are the four components of a time series?

3 What are the guidelines for selecting a forecasting
nique?

4 Selection of the period in moving average method is
erally very scientific. State True or False.

5 Exponential smoothing method is very apt for short-
ı forecasting. State True or False.

5 What techniques could you use to forecast sales if a
product is to be introduced in the market?

istorical analogy b. Delphi method
larket survey. d. Expert opinion
ll of States the above

7 Multiple regression model is suitable for long-term
casting. State True or Flase.

plying the Concepts

8 To measure the effect of advertising and price of the
duct on the demand pattern, the following data were collected form a consumer marketing company for the last 10 months.

Month	Sales (Y) In units	Advertisement Expense (X1) Rs Lacs	Price (X2) In Rs
1	300	45	2400
2	350	50	2200
3	400	55	2100
4	500	85	2000
5	400	65	2150
6	450	60	2000
7	420	57	2150
8	550	68	1950
9	450	60	2050
10	500	70	2000

a. Set up a regression model by taking Sales (Y) as the dependent variable and advertisement expense ($X1$) and price ($X2$) as independent variables. Validate the model using R^2 value.

b. Forecast $X1$ and $X2$ for month 11 by using the exponential smoothing technique.

c. Now forecast sales for month 11 using results obtained in 2).

16.9 The following data refer to the past 12 months sales of a consumer product.

Month	Demand
1	16
2	14
3	12
4	15
5	18
6	21
7	23
8	24
9	25
10	26
11	37
12	38

a. Forecast demand for the 13th month using 3 monthly, 6 monthly, and 9 monthly moving averages.

b. Fit a least square linear trend to the data and project for the month of 13.

c. Use exponential smoothing technique and forecast the demand for month 13. Take $\alpha = 0.3$.

16.10 The following data refer to the sales of commercial vehicles at the All India Level of a leading automobile company in the country during three financial years (April to March).

	Year 1	Year 2	Year 3
April	724	1414	1243
May	1440	2117	1818
June	1606	2199	2880
July	1656	2583	1693
August	1549	2358	2196
Septemebr	2285	3677	3707
October	1523	1823	1931
November	1856	2372	1637
December	2135	2301	1746
January	2119	2761	263
February	2075	2110	265
March	3860	3996	357

a. Draw the time series graph depicting the comparat sales for the three years.

b. Compute 12 monthly moving averages and plot the gr of the moving averages.

c. What is the forecast for April in the fourth year?

16.11 The following data are the sales of a company in past sixteen years. The company is interested in analyz the data in the context of business planning for the n three years. In particular, the company would like to stu the pattern of sales emerging from the data to project sales for the next three years in a reasonable manner.

Year	Sales In Rs. Crores
1	26.20
2	35.40
3	39.20
4	45.80
5	49.00
6	50.40
7	54.80
8	60.00
9	71.80
10	77.60
11	81.40
12	84.60
13	96.80
14	100.60
15	107.60
16	112.00

a. Forecast the sales for the next three years using the le square linear trend line.

b. Comment on the validity of the model by performi appropriate analysis.

REFERENCES

1. Richard I. Levin and David S. Rubin, *Statistics for Management* (Pearson Education, 2004, 7th edn).

2. David R. Anderson, Dennis J. Sweeney, and Thomas Williams, *Statistics for Business and Econom* (Thomson Learning, 2008, 10th edn).

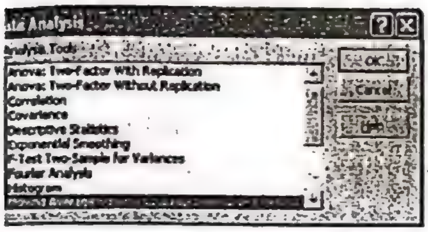
·ENDIX 16

Using Software for Forecasting

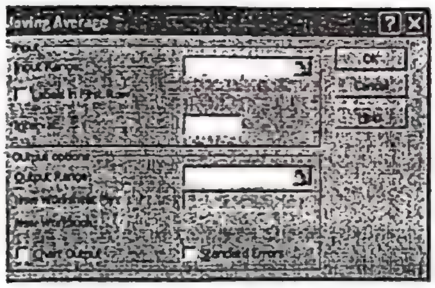
MICROSOFT EXCEL

ımputing Moving Average

ɔls, click Data Analysis, and then click Moving You will get the following screen.



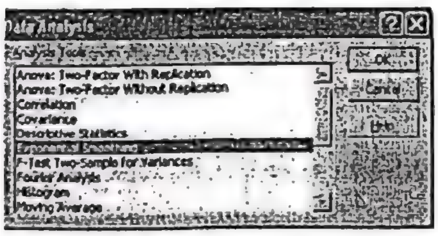
lick OK you will get



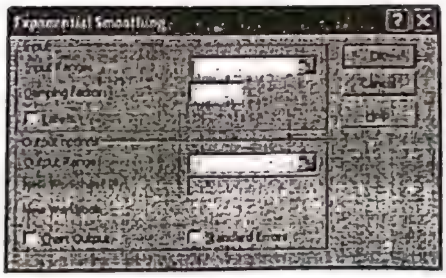
: input range for data, interval and highlight chart :lick OK and you will get the solution.

ımputing Exponential Smoothing

n Microsoft Excel spreadsheet, click *Tools*, click *ılysis*, and click *Exponential Smoothing*. The fol-creen will appear.



Step 2 Click OK and you will get

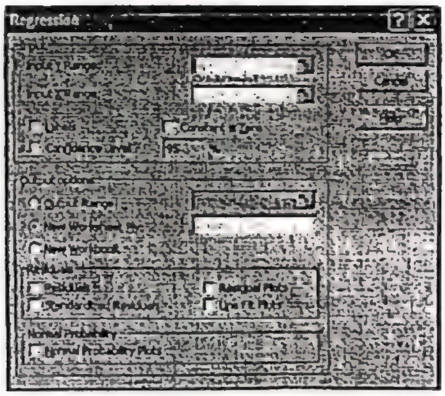


Enter the input range for data, damping factor and highlight chart output. Click OK and you will get the solution.

For Working out Trend Projection Problem

Step 1 Click Tools, click Data-Analysis, and click Regression. You will get

Step 2 Click OK and you will get



Enter the data for *Y* range, *X* range, and click OK. You have the solution now.

CHAPTER 14

Time-Series Analysis

Chapter Contents

Chapter Learning Objectives

When you finish this chapter you should be able to

- Define time-series data and its components.
- Interpret a linear, exponential, or quadratic trend model.
- Fit any common trend model and use it to make forecasts.
- Know the definitions of common fit measures (R^2, *MAPE*, *MAD*, *MSD*).
- Interpret a moving average and use Excel to create it.
- Use exponential smoothing to forecast trendless data.
- Use software to deseasonalize a time-series.
- Use regression with seasonal binaries to make forecasts.

14.1 TIME-SERIES COMPONENTS

Time-Series Data

Businesses must track their performance. By looking at their output over time businesses can tell where they've been, whether they are performing poorly or satisfactorily, and how much improvement is needed, both in the short term and the long term. A ***time-series variable*** (denoted Y) consists of data observed over n periods of time. Consider a clothing retailer that specializes in blue jeans. Examples of time-series data this company might be interested in tracking would be the number of jeans sold and the company's market share. Or, from the manufacturing perspective, the company might track cost of raw materials over time.

VS Chapter 20

Businesses also use time-series data to monitor whether a particular process is stable or unstable. And they use time-series data to help predict the future, a process we call *forecasting*. In addition to business time-series data we see economic time-series data in *The Wall Street Journal* or *BusinessWeek*, and also in *USA Today* or *Time*, or even when we browse the Web. Although business and economic time-series data are most common, we can see time-series data for population, health, crime, sports, and social problems. Usually, time-series data are presented in a graph, like Figures 14.1 and 14.2.

It is customary to plot time-series data either as a line graph or a bar graph, with time on the horizontal axis to reveal how a variable of interest changes over time. In a line graph, the X-Y data points are connected with line segments to make it easier to see fluctuations. While anyone can understand time-series graphs in a general way, this chapter explains how to interpret time-series data *statistically* and to make defensible forecasts. Our analysis begins with sample observations $y_1, y_2, \ldots, y_n$ covering n time periods. The following notation is used:

- y_t is the value of the time-series in period t.
- t is an index denoting the time period ($t = 1, 2, \ldots, n$).
- n is the number of time periods.
- $y_1, y_2, \ldots, y_n$ is the data set for analysis.

FIGURE 14.1

U.S. employment (monthly)

Source: www.clevelandfed.org.

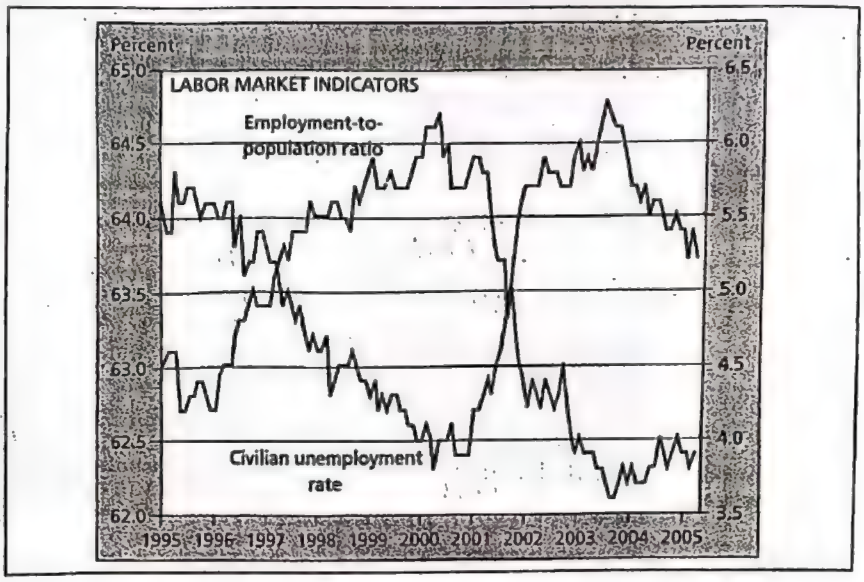
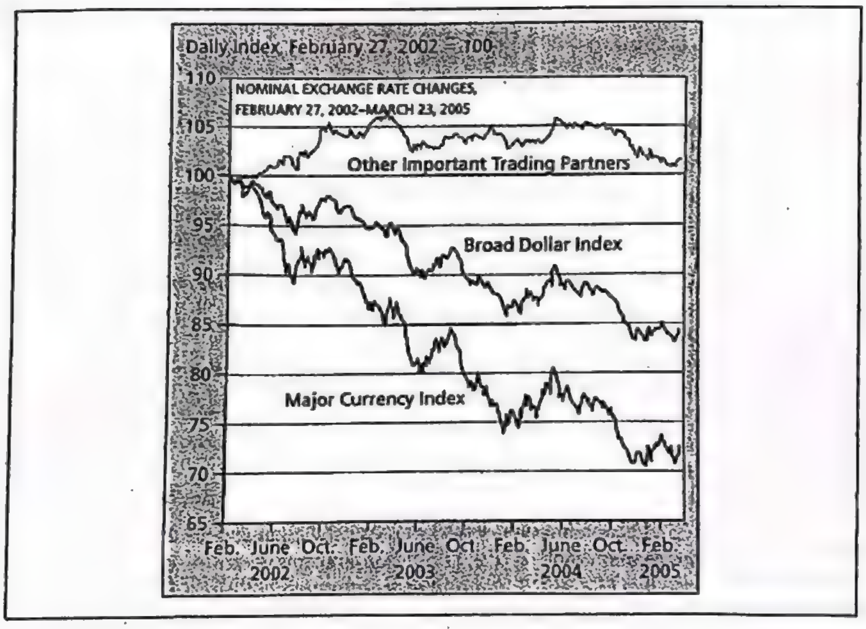


FIGURE 14.2

Exchange rates (daily)

Source: www.clevelandfed.org.



To distinguish time-series data from cross-sectional data, we use y_t instead of x_i for an individual observation, and a subscript t instead of i.

Stocks and Flows

Time-series data may be measured *at a point in time* (a *stock*) or *over an interval of time* (a *flow*). For example, in accounting, balance sheet data are stocks (e.g., measured at the end of the fiscal year) while income statement data are flows (e.g., measured over an entire fiscal year). The concept is quite general. For example, the Gross Domestic Product (GDP) is a flow of goods and services measured *over an interval of time,* while the prime rate of interest is

measured *at a point in time*. Your GPA is measured *at a point in time* while your weekly pay is measured *over an interval of time*. The distinction is sometimes vague in reported data, but a little thought will usually clarify matters. For example, Canada's 2002 unemployment rate (7.6 percent) would be measured at a point in time (e.g., at year's end) while Canada's 2001 electricity production (566.3 billion kWh) would be measured over the entire year.

Periodicity

The ***periodicity*** is the time interval over which data are collected (decade, year, quarter, month, week, day, hour). For example, the U.S. population is measured each *decade*, your personal income tax is calculated *annually*, GDP is reported *quarterly*, the unemployment rate is estimated *monthly*, and *The Wall Street Journal* reports the closing price of General Motors stock *daily* (although stock prices are also monitored continuously on the Web). Firms typically report profits by quarter, but pension liabilities only at year's end. Any periodicity is possible, but the principles of time-series modeling can be understood with these three common data types:

- Annual data (1 observation per year)
- Quarterly data (4 observations per year)
- Monthly data (12 observations per year)

Additive versus Multiplicative Models

Time-series ***decomposition*** seeks to separate a time-series Y into four components: trend (T), cycle (C), seasonal (S), and irregular (I). These components are assumed to follow either an additive or a multiplicative model, as shown in Table 14.1.

TABLE 14.1
Components of a Time-Series

Model	*Components*	*Used For*
Additive	$Y = T + C + S + I$	Data of similar magnitude (short-run or trend-free data) with constant *absolute* growth or decline.
Multiplicative	$Y = T \times C \times S \times I$	Data of increasing or decreasing magnitude (long-run or trended data) with constant *percent* growth or decline.

The additive form is attractive for its simplicity, but the multiplicative model is often more useful for forecasting financial data, particularly when the data vary over a range of magnitudes. Especially in the short run, it may not matter greatly which form is assumed. In fact, the model forms are fundamentally equivalent since the multiplicative model becomes additive if logarithms are taken (as long as the data are nonnegative):

$$\log(Y) = \log(T \times C \times S \times I) = \log(T) + \log(C) + \log(S) + \log(I)$$

A Graphical View

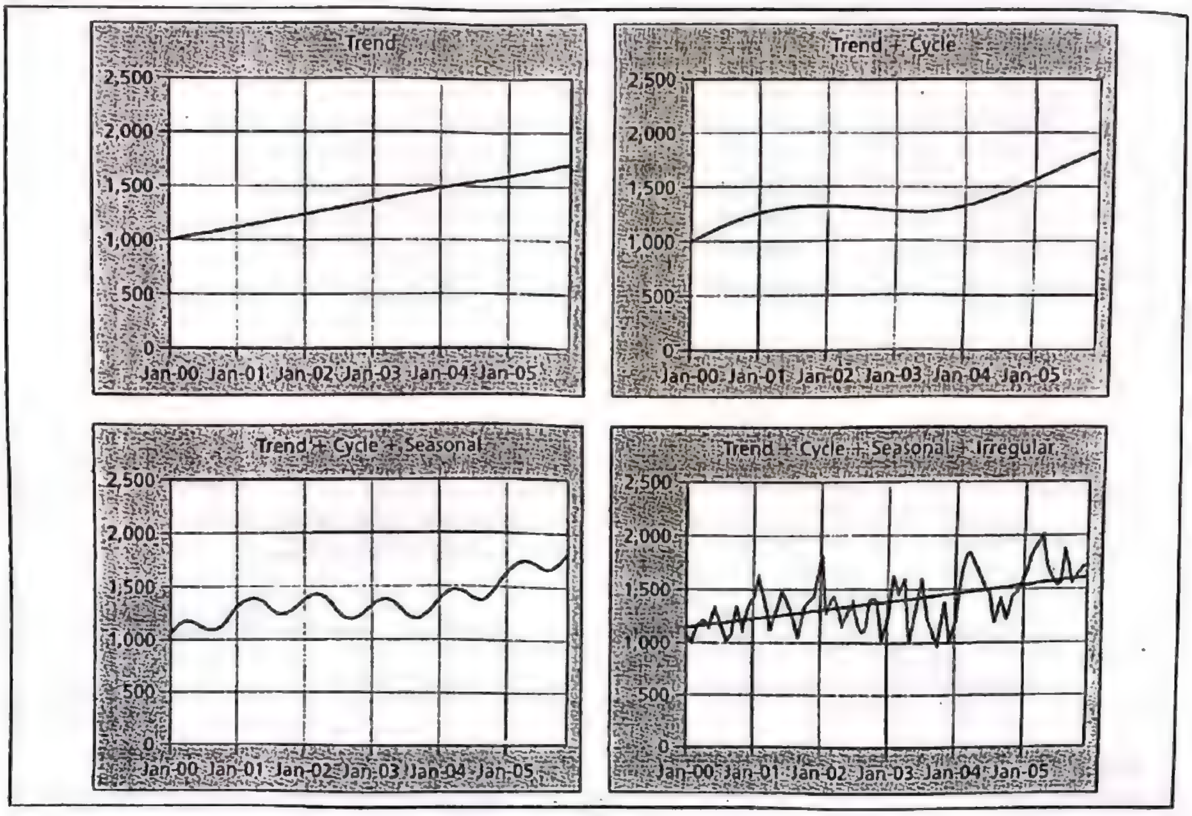
Figure 14.3 illustrates these four components in a hypothetical time series. The four components may be thought of as layering atop one another to produce the actual time series. In this example, the irregular component (I) is large enough to obscure the cycle (C) and seasonal (S) components, but not the trend (T). However, we can usually extract the original components from the time series by using statistical methods.

Trend

Trend (T) is a general movement over all years ($t = 1, 2, \ldots, n$). Change over a few years is not a trend. Some trends are steady and predictable. For example, the data may be steadily growing (e.g., total U.S. population), neither growing nor declining (e.g., your current car's mpg), or steadily declining (infant mortality rates in a developing nation). Most of us think of three general patterns: growth, stability, or decline. But there are subtler trends within each

FIGURE 14.3

Four components of a time series



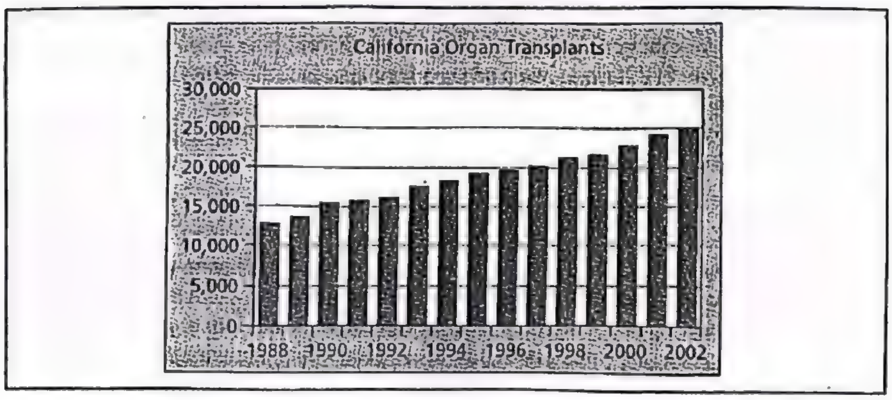
category. A time series can increase at a steady *linear* rate (e.g., the number of books you have read in your lifetime), at an *increasing* rate (e.g., Medicare costs for an aging population), or at a *decreasing* rate (e.g., live attendance at NFL football games). It can grow for awhile and then level off (e.g., sales of HDTV) or grow toward an asymptote (e.g., percent of adults owning a camera phone). A mathematical trend can be fitted to any data, but its predictive value depends on the situation. For example, to predict future organ transplants (Figure 14.4) a mathematical trend might be useful, but a mathematical model might not be very helpful for predicting space launches (Figure 14.5).

FIGURE 14.4

Steady trend

Transplants

Source: www.gsds.org.
© Golden State Donor Services



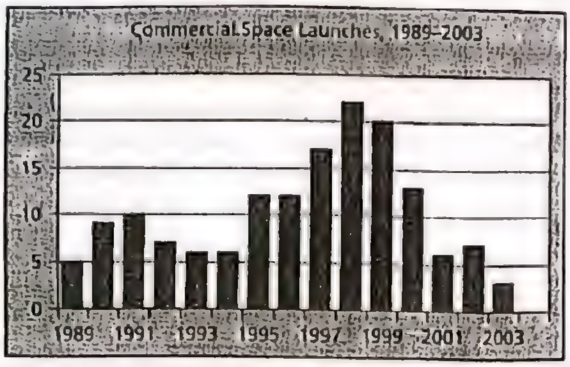


FIGURE 14.5

Erratic pattern

SpaceLaunch

Source: http://ast.faa.gov.

Cycle

Cycle (*C*) is a repetitive up-and-down movement around the trend that covers *several years.* For example, industry analysts have studied cycles for sales of new automobiles, new home construction, inventories, and business investment. These cycles are based primarily on product life and replacement cycles. In any market economy there are broad business cycles that affect employment and production. But there is no general theory of cycles, and even those cycles that have been identified in specific industries have erratic timing and complex causes that defy generalization. Over a small number of time periods (a typical forecasting situation) cycles are undetectable or may resemble a trend. For this reason cycles are not discussed further in this chapter.

Seasonal

Seasonal (*S*) is a repetitive cyclical pattern *within a year.** For example, many retail businesses experience strong sales during the fourth quarter because of Christmas. Automobile sales rise when new models are released. Peak demand for airline flights to Europe occurs during summer vacation travel. Although often imagined as sine waves, seasonal patterns may not be smooth. Peaks and valleys can occur in any month or quarter, and each industry may face its own unique seasonal pattern. For example, June weddings tend to create a "spike" in bridal sales, but there is no "sine wave" pattern in bridal sales. By definition, annual data have no seasonality.

Irregular

Irregular (*I*) is a random disturbance that follows no apparent pattern. It is also called the *error* component or *random noise* reflecting all factors other than trend, cycle, and seasonality. Large error components are not unusual. For example, daily prices of many common stocks fluctuate greatly. When the irregular component is large, it may be difficult to isolate other individual model components. Some data are pure *I* (lacking meaningful *T* or *S* or *C* components). In such cases, we use special techniques (e.g., *moving average* or *exponential smoothing*) to make short-run forecasts. Faced with erratic data, experts may use their own knowledge of a particular industry to make *judgment forecasts.* For example, monthly sales forecasts of a particular automobile may combine judgment forecasts from dealers, financial staff, and economists.

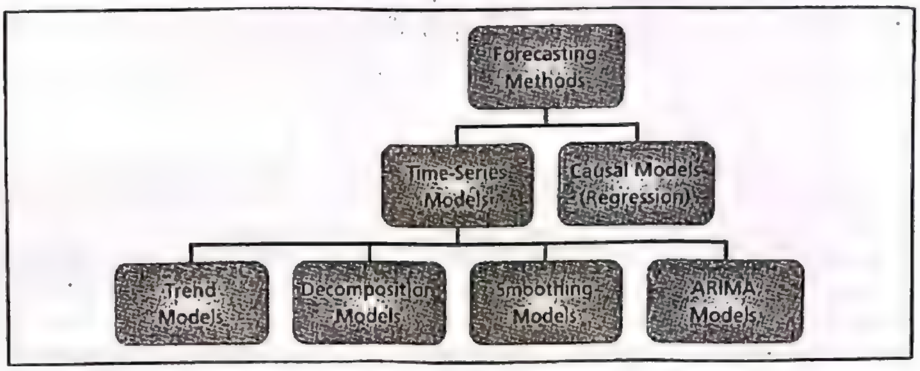
*Repetitive patterns within a week, day, or other time period may also be considered seasonal. For example, mail volume in the U.S. Post Service is higher on Monday. Emergency arrivals at hospitals are lower during the first shift (midnight and 6:00 A.M.). In this chapter, we will discuss only *monthly* and *quarterly* seasonal patterns, because these are most typical of business data.

14.2 TREND FORECASTING

There are many forecasting methods designed for specific situations. Much of this chapter deals with *trend models* because they are so common in business. You will also learn to use *decomposition* to make adjustments for *seasonality*, and how to use *smoothing models*. The important topics of *ARIMA models* and *causal models* using regression are reserved for a more specialized class in forecasting. Figure 14.6 summarizes the main categories of forecasting models.

FIGURE 14.6
Overview of forecasting

VS
Chapter 20



Three Trend Models

There are many possible trend models, but three of them are especially useful in business:

(14.1) $y_t = a + bt$ for $t = 1, 2, \ldots, n$ (linear trend)

(14.2) $y_t = ae^{bt}$ for $t = 1, 2, \ldots, n$ (exponential trend)

(14.3) $y_t = a + bt + ct^2$ for $t = 1, 2, \ldots, n$ (quadratic trend)

The linear and exponential models are widely used because they have only two parameters and are familiar to most business audiences. The quadratic model may be useful when the data have a turning point. All three can be fitted by Excel, MegaStat, or MINITAB. Each model will be examined in turn.

Linear Trend Model

The ***linear trend*** model has the form $y_t = a + bt$. It is useful for a time-series that grows or declines by the same amount (b) in each period, as shown in Figure 14.7. It is the simplest model and may suffice for short-run forecasting. It is generally preferred in business as a baseline forecasting model unless there are compelling reasons to consider a more complex model.

FIGURE 14.7
Linear trend models

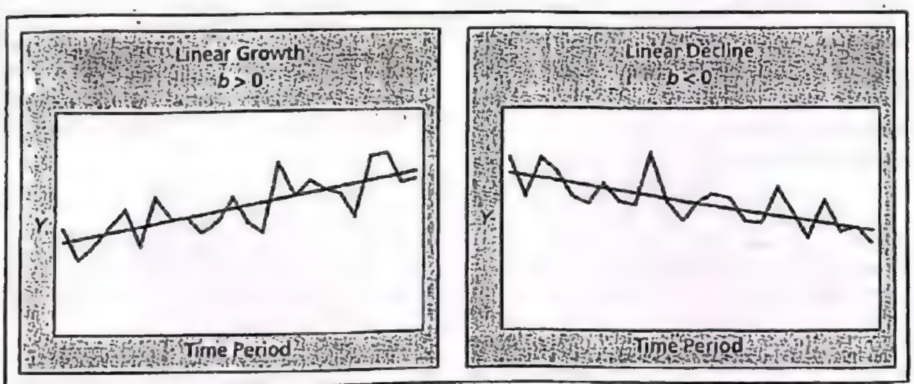


Illustration: Linear Trend

In recent years, the number of U.S. franchised new car dealerships has been declining, due to phasing out of low-volume dealerships and consolidation of market areas. What has been the average annual decline? Based on the line graph in Figure 14.8, the linear model seems appropriate to describe this trend. The slope of Excel's fitted trend indicates that, on average, 235 dealerships are being lost annually.

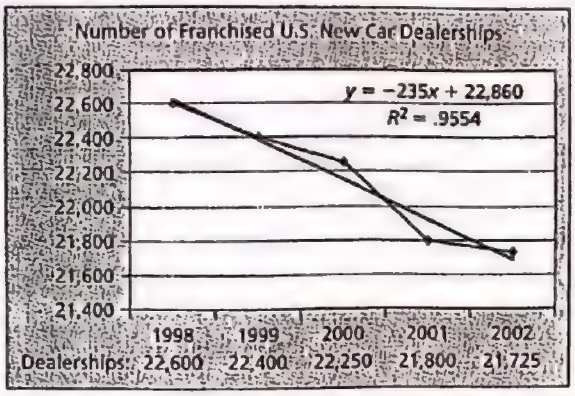


FIGURE 14.8

Excel's Linear trend

CarDealers

Source: *Statistical Abstract of the United States*, 2003.

Linear Trend Calculations

The linear trend is fitted in the usual way by using the ordinary least squares formulas, as illustrated in Table 14.2. Since you are already familiar with regression, we will only point out the use of the index $x_t = 1, 2, 3, 4, 5$ for the calculations (instead of using the years 1998, 1999, 2000, 2001, 2002). We use this time index to simplify the calculations and keep the data magnitudes under control (Excel uses this method too).

Slope: $$b = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{-2,350}{10} = -235$$

Intercept: $$a = \bar{y} - b\bar{x} = 22,155 - (-235)(3) = 22,860$$

Year	x_t	y_t	$x_t - \bar{x}$	$y_t - \bar{y}$	$(x_t - \bar{x})^2$	$(x_t - \bar{x})(y_t - \bar{y})$
1998	1	22,600	−2	445	4	−890
1999	2	22,400	−1	245	1	−245
2000	3	22,250	0	95	0	0
2001	4	21,800	1	−355	1	−355
2002	5	21,725	2	−430	4	−860
Sum	15	110,775	0	0	10	−2,350
Mean	3	22,155	0	0	2	−470

TABLE 14.2
Sums for Least Squares Calculations

Interpreting a Linear Trend

The *slope* of the fitted trend $y_t = 22,860 - 235t$ says that we expect to lose 235 dealerships each year ($dy_t/dt = -235$). The *intercept* is the "starting point" for the time-series in period $t = 0$; that is, $y_0 = 22,860 - 235(0) = 22,860$.

Forecasting a Linear Trend

We can make a forecast for any future year by using the fitted model. In the car dealer example, the fitted trend equation is based on only 5 years' data, so we should be wary of extrapolating very far ahead:

For 2003 ($t = 6$): $y_6 = 22{,}860 - 235(6) = 21{,}450$
For 2004 ($t = 7$): $y_7 = 22{,}860 - 235(7) = 21{,}215$
For 2005 ($t = 8$): $y_8 = 22{,}860 - 235(8) = 20{,}980$

Linear Trend: Calculating R^2

The worksheet shown in Table 14.3 shows the calculation of the coefficient of determination. In this illustration, the linear model gives a good fit ($R^2 = .9554$) to the *past* data. However, a good fit to the past data does not guarantee good *future* forecasts. A deeper analysis of underlying causes of dealership consolidation is needed. What is causing the trend? Are the causal forces likely to remain the same in subsequent years? Could the current trend continue indefinitely, or will it approach an asymptote or limit of some kind? These are questions that forecasters must ask. The forecast is simply a projection of current trend assuming that nothing changes.

$$\textit{Coefficient of determination:}\quad R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = 1 - \frac{25{,}750}{578{,}000} = .9554$$

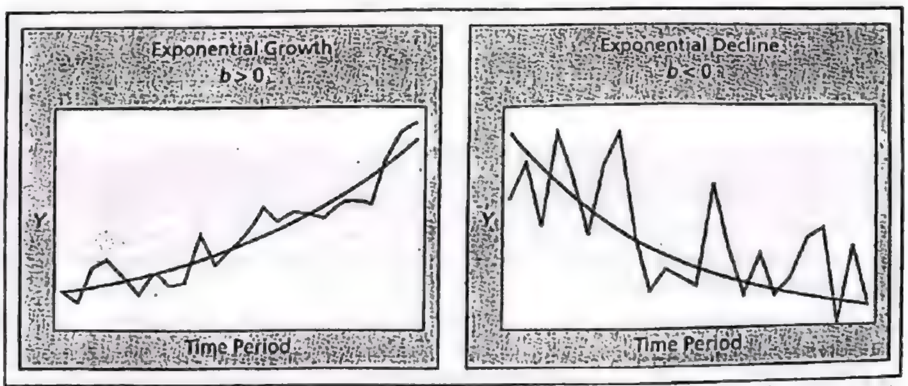
TABLE 14.3
Sums for R^2 Calculations

Year	t	y_t	$\hat{y}_t = 22{,}860 - 235t$	$y_t - \hat{y}_t$	$(y_t - \hat{y}_t)^2$	$(y_t - \bar{y})^2$
1998	1	22,600	22,625	−25	625	198,025
1999	2	22,400	22,390	10	100	60,025
2000	3	22,250	22,155	95	9,025	9,025
2001	4	21,800	21,920	−120	14,400	126,025
2002	5	21,725	21,685	40	1,600	184,900
Sum	15		110,775	0	25,750	578,000

Exponential Trend Model

The ***exponential trend*** model has the form $y_t = ae^{bt}$. It is useful for a time-series that grows or declines at the same *rate* (b) in each period, as shown in Figure 14.9. When the growth rate is positive ($b > 0$), then Y grows by an *increasing* amount each period (unlike the linear model, which assumes a *constant* increment each period). If the growth rate is negative ($b < 0$), then Y declines by a *decreasing* amount each period (unlike the linear model, which assumes a *constant* decrement each period).

FIGURE 14.9
Exponential trend models



When to Use the Exponential Model

The exponential model is often preferred for financial data or data that covers a longer period of time. When you invest money in a commercial bank savings account, interest accrues at a given percent. Your savings grow faster than a linear rate because you earn interest on the accumulated interest. Banks use the exponential formula to calculate interest on CDs. Financial analysts often find the exponential model attractive because costs, revenue, and salaries are best projected under assumed *percent* growth rates.

Another nice feature of the exponential model is that you can compare two growth rates in two time-series variables with dissimilar data units (i.e., a percent growth rate is *unit-free*). For example, between 1990 and 2000 the number of Medicare enrollees grew from 34.3 million persons to 39.6 million persons (1.45 percent growth per annum), while Medicare payments to hospitals grew from \$65.7 billion to \$126.0 billion (6.73 percent growth per annum). Comparing the percents, we see that Medicare insurance payments have been growing more than four times as fast as the Medicare head count. These facts underlie the ongoing debate about Medicare spending in the United States.

There may not be much difference between a linear and exponential model when the growth rate is small and the data set covers only a few time periods. For example, suppose your starting salary is \$50,000. Table 14.4 compares salary increases of \$2,500 each year ($y_t = 50{,}000 + 2{,}500t$) with a continuously compounded 4.879 percent salary growth ($y_t = 50{,}000e^{.04879t}$). Over the first few years, there is little difference. But after 20 years, the difference is obvious, as shown in Figure 14.10. Despite its attractive simplicity* the linear model's assumptions may be inappropriate for some financial variables.

TABLE 14.4
Two Models of Salary Growth

t	$y_t = 50{,}000 + 2{,}500t$ *Linear*	$y_t = 50{,}000\,e^{.04879t}$ *Exponential*
0	50,000	50,000
5	62,500	63,814
10	75,000	81,445
15	87,500	103,946
20	100,000	132,665

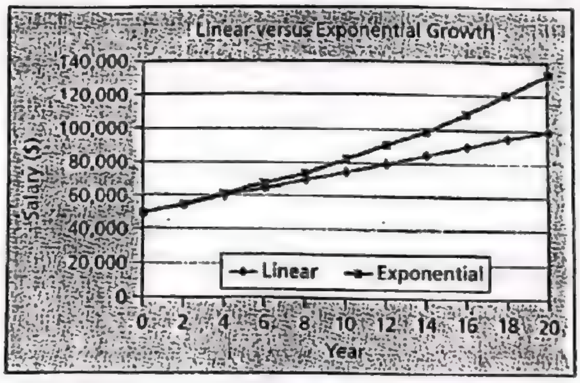


FIGURE 14.10
Linear and exponential growth compared

Illustration: Exponential Trend

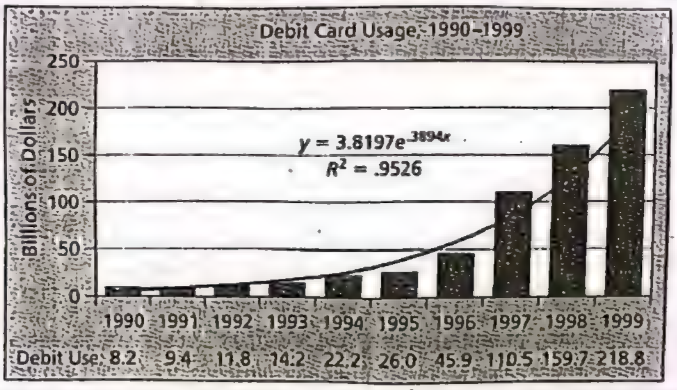
Debit card usage in the United States has shown explosive growth, as indicated in Figure 14.11. Clearly, a linear trend (constant *dollar* growth) would be inadequate. It is more reasonable to assume a constant *percent* rate of growth and fit an exponential model. For the debit

*In a sense, the linear model ($y_t = a + bt$) and the exponential model ($y_t = ae^{bt}$) are equally simple because they are two-parameter models, and a log-transformed exponential model $\ln(y_t) = \ln(a) + bt$ is actually linear.

FIGURE 14.11

Excel's exponential trend

DebitCards



card data, the fitted exponential trend is $y_t = 3.8197e^{.3894t}$. The value of b in the exponential model $y_t = ae^{bt}$ is the continuously compounded growth rate, so we can say that debit card usage is growing at an astonishing rate of 38.94 percent per year. A negative value of b in the equation $y_t = ae^{bt}$ would indicate *decline* instead of growth. The intercept a is the "starting point" in period $t = 0$. For example, $y_0 = 3.8197e^{.3894(0)} = 3.8197$.

Exponential Trend Calculations

Table 14.5 shows the worksheet for the required sums. Calculations of the exponential trend are done by using a transformed variable $z_t = \ln(y_t)$ instead of y_t, to produce a linear equation so that we can use the least squares formulas.

$$\textit{Slope:} \quad b = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(z_t - \bar{z})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{32.12329}{82.5} = .3893732$$

$$\textit{Intercept:} \quad a = \bar{z} - b\bar{x} = 3.481731 - (.3893732)(5.5) = 1.340178$$

When the least squares calculations are completed, we must transform the intercept back to the original units by exponentiation to get the correct intercept $a = e^{1.340178} = 3.8197$. In final form, the fitted trend equation is

$$y_t = ae^{bt} = 3.8197e^{.38937t}$$

TABLE 14.5 Least Squares Sums for the Exponential Model

Year	x_t	y_t	$z_t = \ln(y_t)$	$x_t - \bar{x}$	$z_t - \bar{z}$	$(x_t - \bar{x})^2$	$(x_t - \bar{x})(z_t - \bar{z})$
1990	1	8.2	2.10413	−4.5	−1.37760	20.25	6.19919
1991	2	9.4	2.24071	−3.5	−1.24102	12.25	4.34357
1992	3	11.8	2.46810	−2.5	−1.01363	6.25	2.53408
1993	4	14.2	2.65324	−1.5	−0.82849	2.25	1.24273
1994	5	22.2	3.10009	−0.5	−0.38164	0.25	0.19082
1995	6	26.0	3.25810	0.5	−0.22363	0.25	−0.11182
1996	7	45.9	3.82647	1.5	0.34473	2.25	0.51710
1997	8	110.5	4.70502	2.5	1.22328	6.25	3.05821
1998	9	159.7	5.07330	3.5	1.59157	12.25	5.57048
1999	10	218.8	5.38816	4.5	1.90643	20.25	8.57892
Sum	55	626.7	34.81731	0.0	0.00000	82.5	32.12329
Mean	5.5	62.67	3.481731				

Forecasting an Exponential Trend

We can make a forecast of debit card usage for any future year by using the fitted model*:

For 2001 ($t = 11$): $y_{11} = 3.8197e^{.38937(11)} = 276.8$

For 2002 ($t = 12$): $y_{12} = 3.8197e^{.38937(12)} = 408.5$

For 2003 ($t = 13$): $y_{13} = 3.8197e^{.38937(13)} = 603.0$

Can debit card usage actually continue to grow at a rate of 38.937 percent? It seems unlikely. Typically, when a new product is introduced, its growth rate at first is very strong, but eventually slows down.

Exponential Trend: Calculating R^2

We can calculate R^2 using a worksheet like Table 14.6. Note that all calculations of R^2 are done in terms of $\ln(y_t)$. In this example, the exponential trend gives a very good fit ($R^2 = .9526$) to the past data. Although a high R^2 does not guarantee good forecasts, in the case of debit card usage we might expect the near future to resemble the recent past. Debit cards appear poised to reach a much wider audience of consumers who have traditionally relied on checks or credit cards.

Coefficient of determination: $$R^2 = 1 - \frac{\sum_{t=1}^{n}(z_t - \hat{z}_t)^2}{\sum_{t=1}^{n}(z_t - \bar{z})^2} = 1 - \frac{0.622275}{13.130224} = .9526$$

TABLE 14.6 Sums for R^2 Calculations (Exponential Model)

x_t	$z_t = \ln(y_t)$	$\hat{z}_t = 1.340178 + .389373\,x_t$	$z_t - \hat{z}_t$	$(z_t - \hat{z}_t)^2$	$(z_t - \bar{z})^2$
1	2.10413	1.72955	0.37458	0.14031	1.89777
2	2.24071	2.11892	0.12178	0.01483	1.54013
3	2.46810	2.50830	−0.04020	0.00162	1.02745
4	2.65324	2.89767	−0.24443	0.05975	0.68639
5	3.10009	3.28704	−0.18695	0.03495	0.14565
6	3.25810	3.67642	−0.41832	0.17499	0.05001
7	3.82647	4.06579	−0.23933	0.05728	0.11884
8	4.70502	4.45516	0.24985	0.06243	1.49643
9	5.07330	4.84454	0.22876	0.05233	2.53308
10	5.38816	5.23391	0.15425	0.02379	3.63446
Sum	34.81731	34.81731	0	0.622275	13.130224
Mean	3.48173				

Quadratic Trend Model

The *quadratic trend* model has the form $y_t = a + bt + ct^2$. It is useful for a time series that has a turning point or that is not captured by the exponential model. If $c = 0$, the quadratic model $y_t = a + bt + ct^2$ becomes a linear model because the term ct^2 drops out of the

*Excel uses the exponential formula $y_t = ae^{bt}$ in which the coefficient b is the *continuously compounded* growth rate. But MINITAB uses an equivalent formula $y_t = y_0(1 + r)^t$, which you may recognize as the formula for compound interest. Although the formulas appear different, they give identical forecasts. For example, for the debit card data, MINITAB's fitted trend is $y_t = 3.81972(1.47606)^t$ so the forecasts are

For 2001 ($t = 11$): $y_{11} = 3.81972(1.47606)^{11} = 276.8$

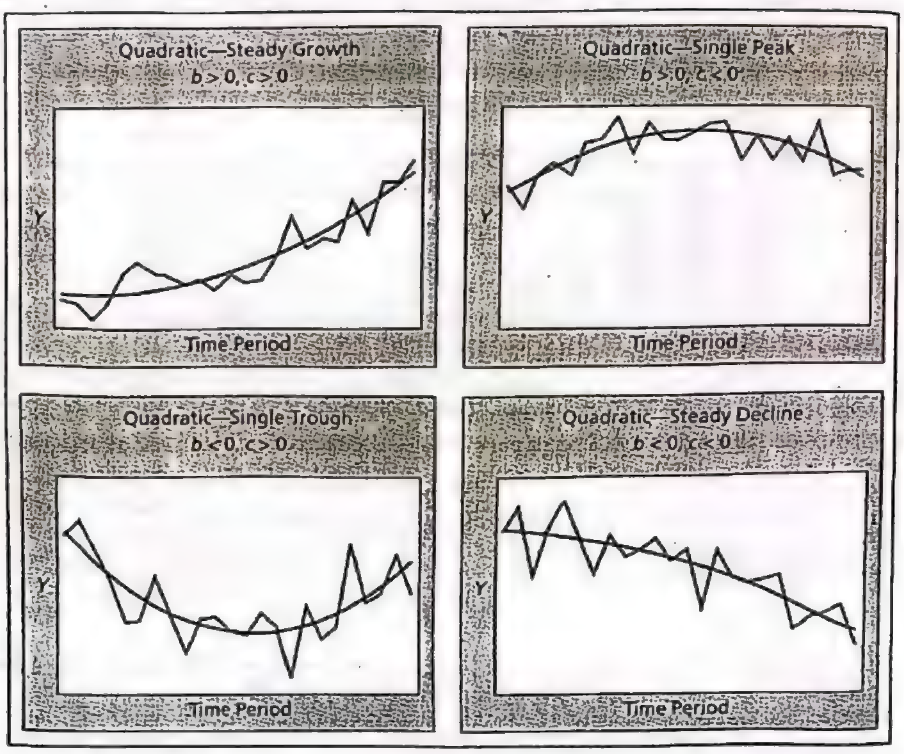
For 2002 ($t = 12$): $y_{12} = 3.81972(1.47606)^{12} = 408.5$

For 2003 ($t = 13$): $y_{13} = 3.81972(1.47606)^{13} = 603.0$

To convert MINITAB's fitted equation to Excel's, set $a = y_0$ and $b = \ln(1 + r)$. To convert Excel's fitted equation to MINITAB's, set $y_0 = a$ and $r = e^b - 1$.

FIGURE 14.12

Four quadratic trend models



equation (i.e., the linear model is a special case of the quadratic model). Some forecasters fit a quadratic model as a way of checking for nonlinearity. If the coefficient c does not differ significantly from zero, then the linear model would suffice. Depending on the values of b and c, the quadratic model can assume any of four shapes, as shown in Figure 14.12.

Illustration: Quadratic Trend

The number of hospital beds (Table 14.7) in the United States declined during the late 1990s, but then showed signs of leveling out or even increasing again. What trend would we choose if the objective is to make a realistic 1-year forecast?

TABLE 14.7

U.S. Hospital Beds, 1995–2001

HospitalBeds

Source: *Statistical Abstract of the United States, 2003*

Year	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*
Beds (000)	1,081	1,062	1,035	1,013	994	984	987

Figures 14.13 and 14.14 show 1-year projections using the linear and quadratic models. Many observers would think that the quadratic model offers a more believable prediction, because the quadratic model is able to capture the slight curvature in the data pattern. But this gain in forecast credibility must be weighed against the added complexity of the quadratic model. It appears that the forecasts would turn upward if projected more than 1 year ahead. We should be especially skeptical of any polynomial model that is projected more than one or two periods into the future.

Because the quadratic trend model $y_t = a + bt + ct^2$ is a multiple regression with two predictors (t and t^2), the least squares calculations are not shown. However, Figure 14.15 shows the MINITAB fitted regression. Note that both t and t^2 are significant predictors (large t, small p).

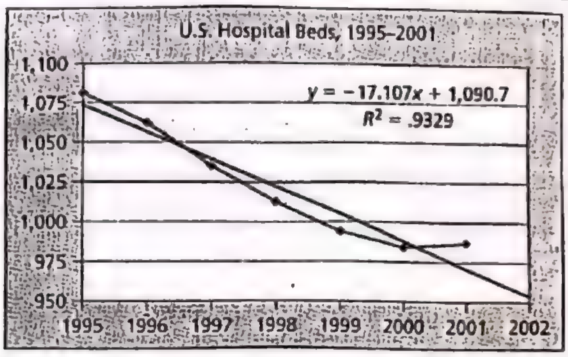


FIGURE 14.13
Linear trend

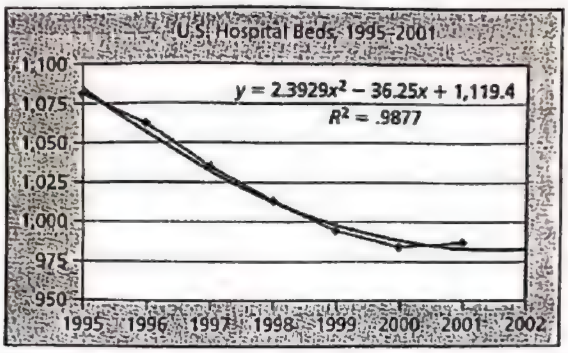


FIGURE 14.14
Quadratic trend

The regression equation is
Beds = 1119 − 36.2 Time + 2.39 Time2

Predictor	Coef	SE Coef	T	P
Constant	1119.43	8.10	138.15	0.000
Time	−36.250	4.644	−7.81	0.001
Time2	2.3929	0.5673	4.22	0.014

S = 5.19959 R-Sq = 98.8% R-Sq(adj) = 98.2%

FIGURE 14.15
MINITAB's quadratic regression

Using Excel for Trend Fitting

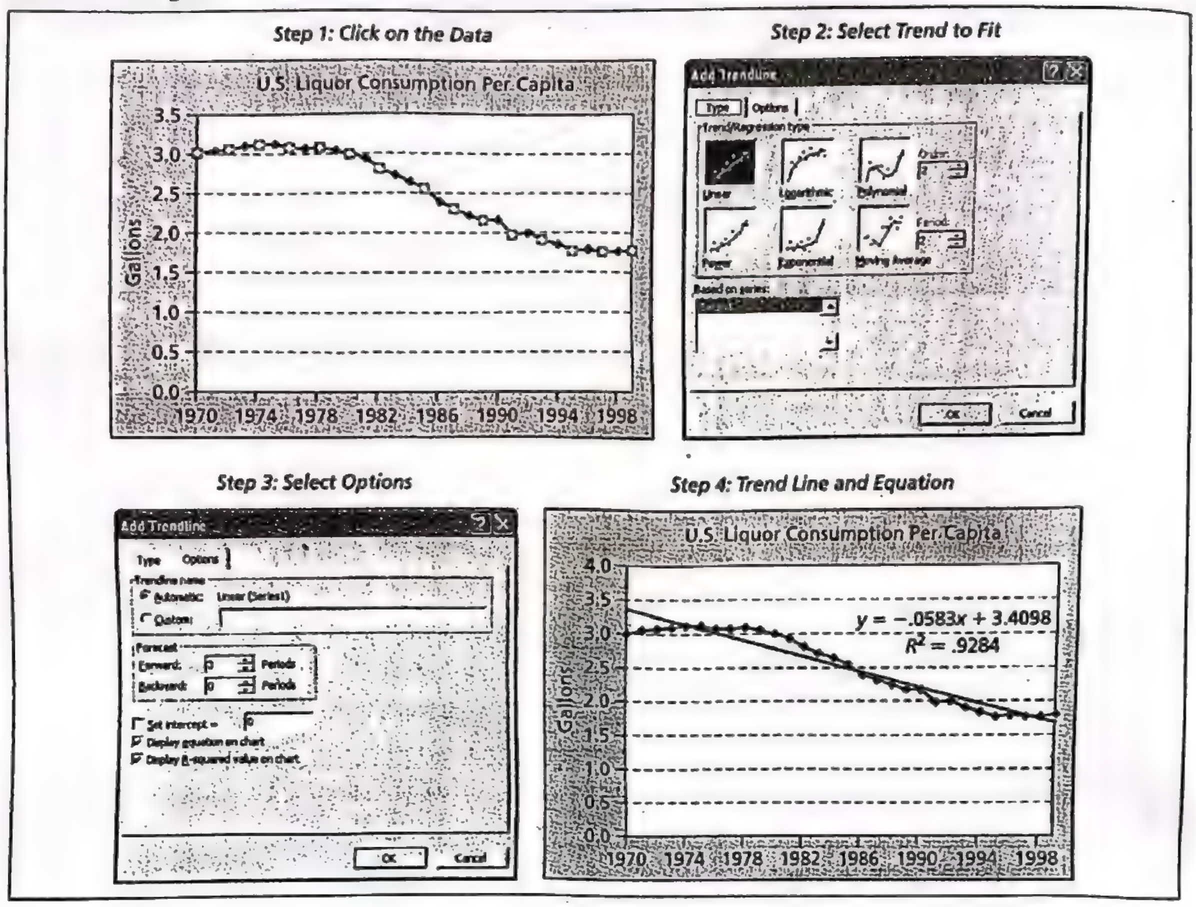
Plot the data, right-click on the data, and choose a trend. Figure 14.16 shows Excel's menu of six trend options. The menu includes a sketch of each trend type. Click the Options tab if you want to display the R^2 and fitted equation on the graph, or if you want to plot forecasts (trend extrapolations) on the graph. The quadratic model is a ***polynomial model*** of order 2. Despite the many choices, some patterns cannot be captured by any of the common trend models. By default, Excel only reports four decimal accuracy. However, you can click on Excel's fitted trend equation, choose Format Data Labels, choose Number, and set the number of decimal places you want to see.

Trend-Fitting Criteria

It is so easy to fit a trend in Excel that it is tempting to "shop around" for the best fit. But forecasters prefer the simplest trend model that adequately matches the trend (the principle of

FIGURE 14.16

Excel's trend-fitting menus



Occam's Razor). Simple models are easier to interpret and explain to others. Criteria for selecting a trend model for forecasting include:

Criterion	*Ask Yourself*
• Occam's Razor	Would a simpler model suffice?
• Overall fit	How does the trend fit the past data?
• Believability	Does the extrapolated trend "look right"?
• Fit to recent data	Does the fitted trend match the last few data points?

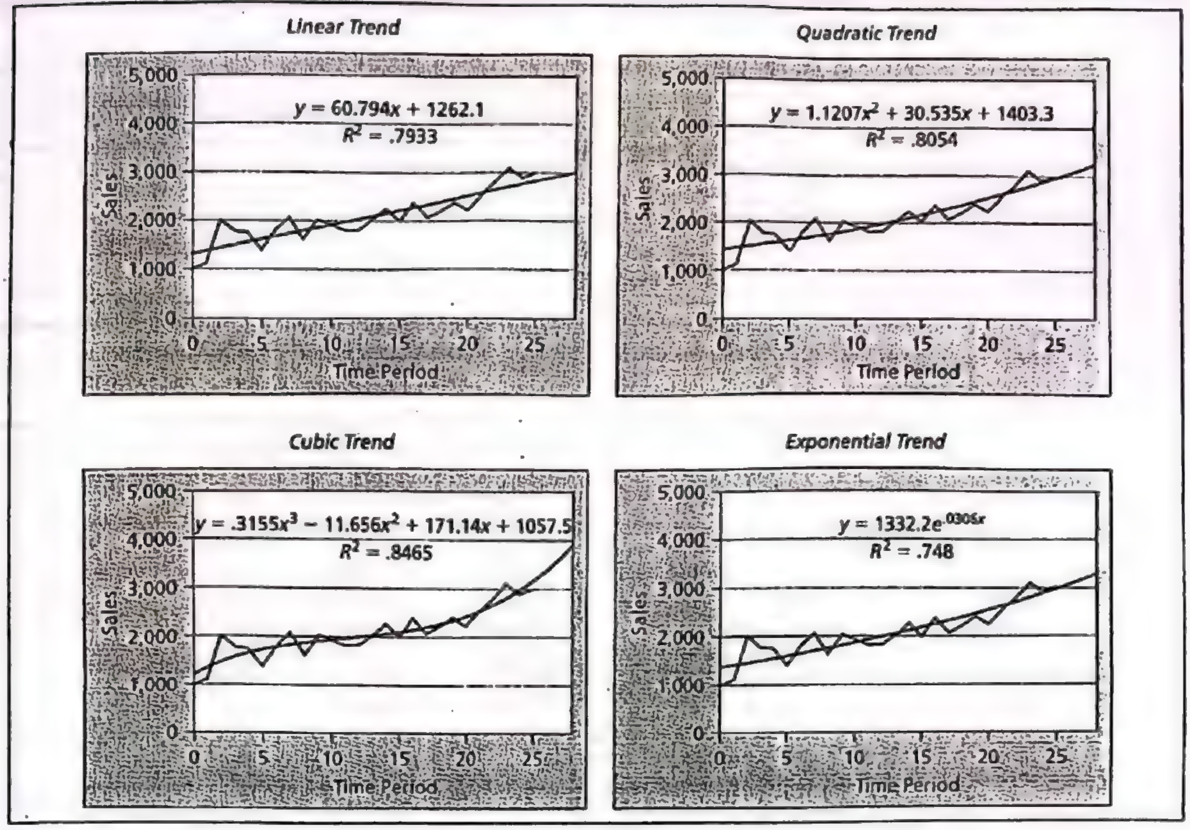
EXAMPLE

Comparing Trends

You can usually increase the R^2 by choosing a more complex model. But if you are making a *forecast,* this is not the only relevant issue, because R^2 measures the fit to the *past* data. Figure 14.17 shows four fitted trends using the same data, with three-period forecasts. For this data set, the linear model may be inadequate because its fit to recent periods is marginal (we prefer the simplest model *only if* it "does the job"). Here, the cubic trend yields the highest R^2, but the fitted equation is nonintuitive and would be hard to explain or defend. Also, its forecasts appear to be increasing too rapidly. In this example, the exponential model has the lowest R^2, yet matches the recent data fairly well and its forecasts appear credible when projected a few periods ahead.

FIGURE 14.17

Four fitted trends using the same data



Any trend model's forecasts become less reliable as they are extrapolated farther into the future. The quadratic trend, the simplest of Excel's polynomial models, is sometimes acceptable for short-term forecasting. However, forecasters avoid higher-order polynomial models (cubic and higher) not only because they are complex, but also because they can give bizarre forecasts when extrapolated more than one period ahead. Table 14.8 compares the features of the three most common trend models.

TABLE 14.8 Comparison of Three Trend Models

Model	*Pro*	*Con*
Linear	1. Simple, familiar to everyone. 2. May suffice for short-run data.	1. Assumes constant slope. 2. Cannot capture nonlinear change.
Exponential	1. Familiar to financial analysts. 2. Shows compound percent growth rate.	1. Some managers are unfamiliar with e^x. 2. Data values must be positive.
Quadratic	1. Useful for data with a turning point. 2. Useful test for nonlinearity.	1. Complex and no intuitive interpretation. 2. Can give untrustworthy forecasts if extrapolated too far.

SECTION EXERCISES

14.1 (a) Make an Excel graph of the data on U.S. diesel new car sales. (b) Discuss the underlying causes that might explain the trend. (c) Use Excel, MegaStat, or MINITAB to fit three trends (linear, quadratic, and exponential) to the time series. (d) Which trend model do you think is best to make forecasts for the next 3 years? Why? (e) Use *each* of the three fitted trend equations to make numerical forecasts for 2004, 2005, and 2006. How much difference does the choice of model make? Which forecasts do you trust the most, and why? (f) If you have access to *Ward's Automotive Yearbook*, check your forecasts. How accurate were they? **Diesel**

U.S. Diesel New Car Sales, 1993–2003

Year	*Sales*
1993	2,800
1994	3,577
1995	3,139
1996	8,469
1997	7,331
1998	10,972
1999	13,573
2000	22,634
2001	15,077
2002	31,430
2003	38,524

Source: *Ward's Automotive Yearbook*, 2004, 66th ed., p. 40.

14.2 (a) Make an Excel graph of the data on U.S. online advertising spending. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Use Excel, MegaStat, or MINITAB to fit three trends (linear, quadratic, exponential) to the time-series. (d) Which trend model do you think is best to make forecasts for the next 3 years? Why? (e) Use *each* of the three fitted trend equations to make a numerical forecast for 2007. How similar are the three models' forecasts? **Online**

U.S. Online Advertising, 2000–2006 (billions)

Year	*Spending*
2000	8.1
2001	7.1
2002	6.0
2003	6.3
2004	6.8
2005	7.2
2006	8.1

Source: William F. Arens, *Contemporary Advertising*, 9th ed. (McGraw-Hill, 2004), p. 549.

14.3 (a) Make an Excel line graph of the data on computer viruses. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Fit three trends (linear, exponential, quadratic). (d) Which trend model is best, and why? If none is satisfactory, explain. (e) Make a forecast for 2003, using a trend model of your choice or a judgment forecast. **PCViruses**

Virus Infections Per Month Per 1,000 PCs

Year	Viruses
1996	10
1997	21
1998	32
1999	80
2000	91
2001	103
2002	105

Source: *PC Magazine* 22 no. 9 (May 27, 2003), p. 23.

14.4 (a) Make an Excel line graph of the work hours data. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Fit three trends (linear, exponential, quadratic). (d) Which trend model is best, and why? If none is satisfactory, explain. (e) Make a forecast for 2000, using a trend model of your choice or a judgment forecast. WorkHours

Average Annual Hours of Work, Married Couple with Children

Year	Hours	Year	Hours
1982	3,160	1991	3,515
1983	3,194	1992	3,491
1984	3,289	1993	3,536
1985	3,309	1994	3,589
1986	3,391	1995	3,616
1987	3,451	1996	3,638
1988	3,460	1997	3,679
1989	3,534	1998	3,685
1990	3,536	1999	3,714

Source: *Statistical Abstract of the United States, 2002.*

14.5 (a) Plot the data on fruit and vegetable consumption. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Fit a linear trend to the data. (d) Interpret the trend equation. What are its implications for producers? (e) Make a forecast for 2005. *Note:* Time increments are 5 years, so use $t = 6$ for your 2005 forecast. Fruits

U.S. Per Capita Consumption of Commercially Produced Fruits and Vegetables (pounds)

Year	Total
1980	608.0
1985	629.3
1990	659.2
1995	690.0
2000	705.4

Source: *Statistical Abstract of the United States, 2003.*

Mini Case 14.1

U.S. Trade Deficit

The imbalance between imports and exports (Table 14.9) has been a vexing policy problem for U.S. policymakers for decades. The last time the United States had a trade surplus was in 1975, partly due to reduced dependency on foreign oil through conservation measures enacted

after the oil crisis (shortages and gas lines) in the early 1970s. However, the trade deficit has become more acute over time, due partly to continued oil imports, and, more recently, to availability of cheaper goods from China and other emerging economies.

TABLE 14.9 U.S. International Trade, 1990–2002 TradeDeficit

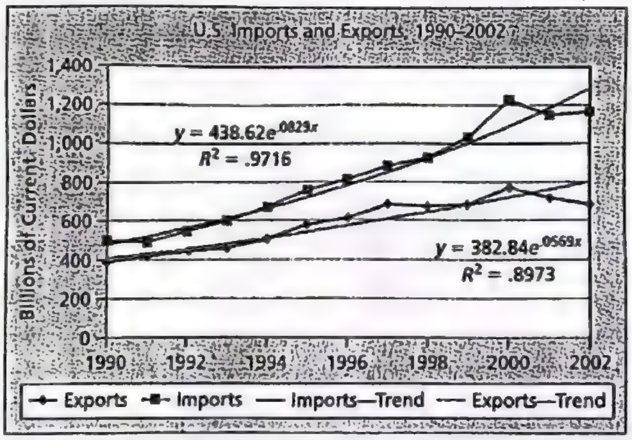
Period	Year	Exports	Imports
1	1990	387.4	498.4
2	1991	414.1	491.0
3	1992	439.6	536.5
4	1993	456.9	589.4
5	1994	502.9	668.7
6	1995	575.2	749.4
7	1996	612.1	803.1
8	1997	678.4	876.5
9	1998	670.4	917.1
10	1999	684.0	1,030.0
11	2000	772.0	1,224.4
12	2001	718.7	1,145.9
13	2002	681.9	1,164.7

Source: *Statistical Abstract of the United States, 2003*. Figures are in billions of current dollars.

Figure 14.18 shows the data graphically, with fitted exponential trends. The trends fit well except for 2001–2002. Imports fell in the recession that began in 2000, but then began to pick up, while exports remained weak. The fitted trend equations reveal that imports have been growing at a compound annual rate of 8.29 percent, while exports have only grown at a compound annual rate of 5.69 percent.

FIGURE 14.18

U.S. trade, 1990–2002



If we project these disparate growth rates, we would predict a widening trade deficit (calculations shown below). Of course, the assumption of *ceteris paribus* may not hold. Policymakers may seek to weaken the dollar or to change fuel efficiency of U.S. vehicles, or there could be changes in foreign economies (e.g., China). Forecasts are less a way of predicting the future than of showing where we are heading *if* nothing changes. A paradox of forecasting is that, as soon as decision makers see the implications of the forecast, they take steps to make sure the forecast is wrong!

Year (Period)	*Imports Trend Projection*	*Exports Trend Projection*
2004 (15)	$y_t = 438.62e^{.0829(15)} = 1{,}521$	$y_t = 382.84e^{.0569(15)} = 899$
2006 (17)	$y_t = 438.62e^{.0829(17)} = 1{,}795$	$y_t = 382.84e^{.0569(17)} = 1{,}007$
2008 (19)	$y_t = 438.62e^{.0829(19)} = 2{,}119$	$y_t = 382.84e^{.0569(19)} = 1{,}129$
2010 (21)	$y_t = 438.62e^{.0829(21)} = 2{,}501$	$y_t = 382.84e^{.0569(21)} = 1{,}265$

14.3 ASSESSING FIT

Five Measures of Fit

In time-series analysis, you are likely to encounter several different measures of "fit" that show how well the estimated trend model matches the observed time series. "Fit" refers to historical data, and you should bear in mind that a good fit is no guarantee of good forecasts—the usual goal. Five common measures of fit are shown in Table 14.10.

TABLE 14.10 Five Measures of Fit

Statistic	*Description*	*Pro*	*Con*
(14.4) $R^2 = 1 - \dfrac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2}$	Coefficient of determination	1. Unit-free measure. 2. Very common.	1. Often interpreted incorrectly (e.g., "percent of correct predictions").
(14.5) $MAPE = \dfrac{100}{n}\sum_{t=1}^{n}\dfrac{\lvert y_t - \hat{y}_t\rvert}{y_t}$	Mean Absolute Percent Error (MAPE)	1. Unit-free measure (%). 2. Intuitive meaning.	1. Requires $y_t > 0$. 2. Lacks nice math properties.
(14.6) $MAD = \dfrac{1}{n}\sum_{t=1}^{n}\lvert y_t - \hat{y}_t\rvert$	Mean Absolute Deviation (MAD)	1. Intuitive meaning. 2. Same units as y_t.	1. Not unit-free. 2. Lacks nice math properties.
(14.7) $MSD = \dfrac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$	Mean Squared Deviation (MSD)	1. Nice math properties. 2. Penalizes big errors more.	1. Nonintuitive meaning. 2. Rarely reported.
(14.8) $SE = \sqrt{\sum_{t=1}^{n}\dfrac{(y_t - \hat{y}_t)^2}{n-2}}$	Standard error	1. Same units as y_t. 2. For confidence intervals.	1. Nonintuitive meaning.

EXAMPLE

Fire Losses

Figure 14.19 shows a MINITAB graph with fitted linear trend and 3-year forecasts for aggregate U.S. fire losses between 1980 and 2000. Notice that, instead of R^2, MINITAB displays *MAPE*, *MAD*, and *MSD* statistics. Table 14.11 shows the calculations for these statistics

FIGURE 14.19

MINITAB's time-series trend—linear model FireLosses

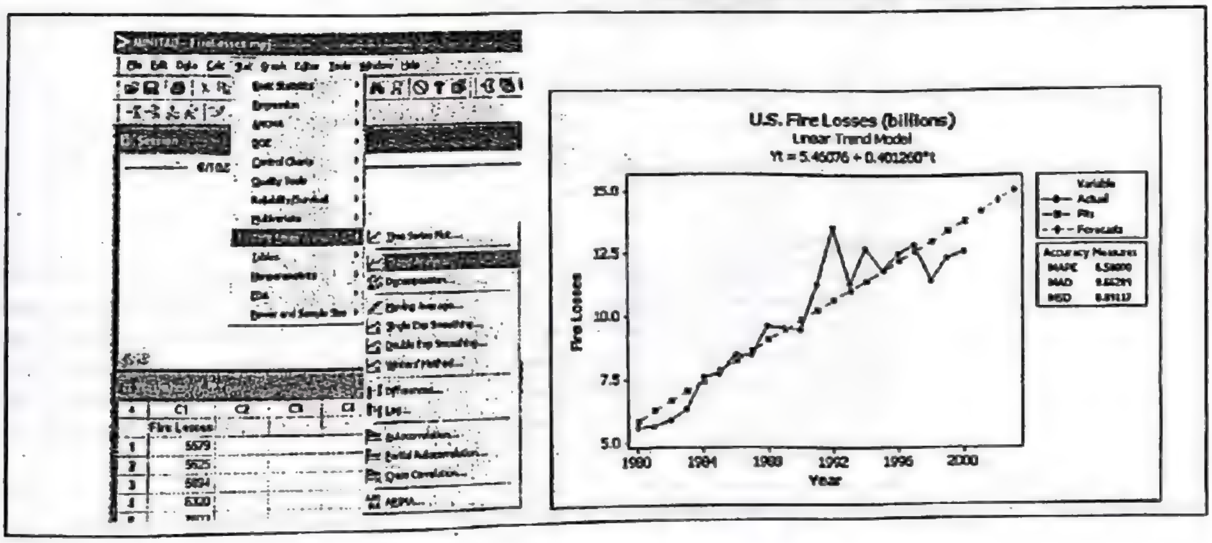


TABLE 14.11 Sums for *MAD*, *MAPE*, *MSD*, and Standard Error FireLosses

Period	Year	Losses	$\hat{y}_t = 5.46076 + .401260t$	$y_t - \hat{y}_t$	$\lvert y_t - \hat{y}_t \rvert$	$\lvert y_t - \hat{y}_t \rvert / y_t$	$(y_t - \hat{y}_t)^2$
1	1980	5.579	5.8620	−0.2830	0.2830	0.0507	0.0801
2	1981	5.625	6.2633	−0.6383	0.6383	0.1135	0.4074
3	1982	5.894	6.6645	−0.7705	0.7705	0.1307	0.5937
4	1983	6.320	7.0658	−0.7458	0.7458	0.1180	0.5562
5	1984	7.602	7.4671	0.1349	0.1349	0.0178	0.0182
6	1985	7.753	7.8683	−0.1153	0.1153	0.0149	0.0133
7	1986	8.488	8.2696	0.2184	0.2184	0.0257	0.0477
8	1987	8.504	8.6708	−0.1668	0.1668	0.0196	0.0278
9	1988	9.626	9.0721	0.5539	0.5539	0.0575	0.3068
10	1989	9.514	9.4734	0.0406	0.0406	0.0043	0.0017
11	1990	9.495	9.8746	−0.3796	0.3796	0.0400	0.1441
12	1991	11.302	10.2759	1.0261	1.0261	0.0908	1.0529
13	1992	13.588	10.6771	2.9109	2.9109	0.2142	8.4731
14	1993	11.331	11.0784	0.2526	0.2526	0.0223	0.0638
15	1994	12.778	11.4797	1.2983	1.2983	0.1016	1.6857
16	1995	11.887	11.8809	0.0061	0.0061	0.0005	0.0000
17	1996	12.544	12.2822	0.2618	0.2618	0.0209	0.0686
18	1997	12.940	12.6834	0.2566	0.2566	0.0198	0.0658
19	1998	11.510	13.0847	−1.5747	1.5747	0.1368	2.4797
20	1999	12.428	13.4860	−1.0580	1.0580	0.0851	1.1193
21	2000	12.659	13.8872	−1.2282	1.2282	0.0970	1.5085
			Sum	0.00	13.9206	1.3818	18.7145
			Mean	0.00	0.66289	0.0658	0.89117

of fit. Since the residuals $y_t - \hat{y}_t$ sum to zero, we see why it's necessary to sum either their absolute values or their squares to obtain a measure of fit. *MAPE*, *MAD*, *MSD* and SE would be zero if the trend provided a perfect fit to the time series.

Calculations

Using the sums in Table 14.11, we can apply the formulas for each fit statistic:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t} = \frac{100}{21}(1.3818) = 6.58\%$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| = \frac{1}{21}(13.9206) = .66289$$

$$MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = \frac{1}{21}(18.7145) = .89117$$

$$SE = \sqrt{\sum_{t=1}^{n}\frac{(y_t - \hat{y}_t)^2}{n-2}} = \sqrt{\frac{18.7145}{21-2}} = .99246$$

Interpretation

The *MAPE* says that our fitted trend has a mean absolute error of 6.58 percent. The *MAD* says that the average error is .66289 billion dollars (ignoring the sign). The *MSD* lacks a simple interpretation. These fit statistics are most useful in comparing different trend models for the same data. All the statistics (especially the *MSD*) are affected by the unusual residual in 1992, when fire losses greatly exceeded the trend. The standard error is useful if we want to make a prediction interval for a forecast, using formula 14.9. It is the same formula you saw in Chapter 12.

(14.9) $$\hat{y}_t \pm t_{n-2}\text{SE}\sqrt{1 + \frac{1}{n} + \frac{(x_t - \bar{x})^2}{\sum_{t=1}^{n}(x_t - \bar{x})^2}}$$ (prediction interval for future y_t)

You may recall from Chapter 12 that you can get a "quick" approximate 95 percent prediction interval by using $\hat{y}_t \pm 2$ SE. However, for forecasts beyond the range of the observed data, you should use formula 14.9, which widens the confidence intervals when the time index is far from its historic mean.

14.4 MOVING AVERAGES

Trendless or Erratic Data

What if the time series $y_1, y_2, \ldots, y_n$ is erratic or has no consistent trend? In such cases, there may be little point in fitting a trend. A conservative approach is to calculate a *moving average*. There are two main types of moving averages: trailing or centered. We will illustrate each.

Trailing Moving Average (*TMA*)

The simplest kind of moving average is the ***trailing moving average*** (***TMA***) over the last m periods.

$$\hat{y}_t = \frac{y_t + y_{t-1} + \cdots + y_{t-m+1}}{m} \quad \text{(trailing moving average over } m \text{ periods)} \quad (14.10)$$

The *TMA* smoothes the past fluctuations in the time-series, helping us see the pattern more clearly. The choice of m depends on the situation. A larger m yields a "smoother" *TMA*, but requires more data. The value of $\hat{y}_t$ may also be used as a forecast for period $t + 1$. Beyond the range of the observed data $y_1, y_2, \ldots, y_n$ there is no way to update the moving average, so it is best regarded as a *one-period-ahead forecast*.

EXAMPLE

Fuel Economy

Many drivers keep track of their fuel economy. For a given vehicle, there is likely to be little trend over time, but there is always random fluctuation. Also, current driving conditions (e.g., snow, hot weather, road trips) could temporarily affect mileage over several consecutive time periods. In this situation, a moving average might be considered. Table 14.12 shows

TABLE 14.12 Andrew's Miles Per Gallon ($n = 20$) AndrewsMPG

Obs	*Date*	*Miles Driven*	*Gallons*	*MPG*	*TMA*	*CMA*
1	5-Jan	285	11.324	25.168		
2	7-Jan	185	8.731	21.189		23.074
3	11-Jan	250	10.934	22.864	23.074	22.815
4	15-Jan	296	12.135	24.392	22.815	22.905
5	19-Jan	232	10.812	21.458	22.905	23.326
6	25-Jan	301	12.475	24.128	23.326	22.158
7	30-Jan	285	13.645	20.887	22.158	22.581
8	3-Feb	263	11.572	22.727	22.581	22.747
9	7-Feb	250	10.152	24.626	22.747	23.856
10	14-Feb	307	12.678	24.215	23.856	23.283
11	22-Feb	242	11.520	21.007	23.283	22.942
12	29-Feb	288	12.201	23.605	22.942	22.937
13	5-Mar	285	11.778	24.198	22.937	24.103
14	8-Mar	313	12.773	24.505	24.103	22.638
15	13-Mar	283	14.732	19.210	22.638	23.330
16	18-Mar	318	12.103	26.274	23.330	21.620
17	22-Mar	195	10.064	19.376	21.620	23.746
18	28-Mar	320	12.506	25.588	23.746	22.904
19	2-Apr	270	11.369	23.749	22.904	23.910
20	12-Apr	259	11.566	22.393	23.910	

Source: Data were collected by statistics student Andrew Fincher for his 11-year-old Pontiac Bonneville 3.8L V6.

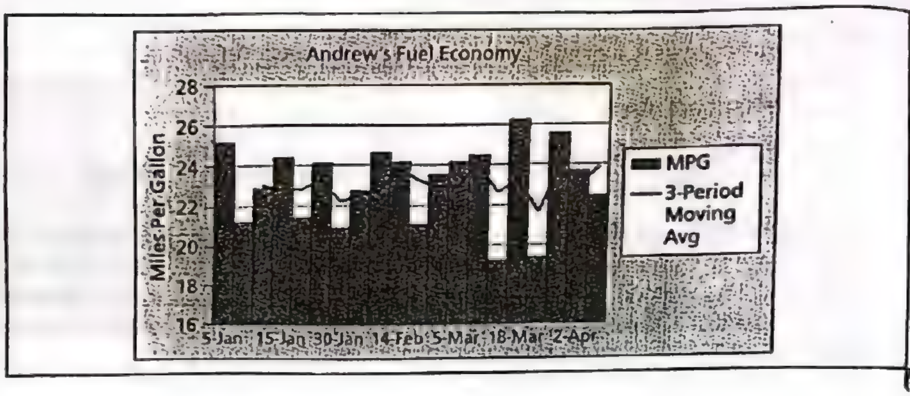
Andrew's fuel economy data set. Column six shows a three-period *TMA*. For example, for period 6 (yellow-shaded cells) the *TMA* is

$$\hat{y}_6 = \frac{24.392 + 21.458 + 24.128}{3} = 23.326$$

It is easiest to appreciate the moving average's "smoothing" of the data when it is displayed on a graph, as in Figure 14.20. It is clear that Andrew's mean is around 23 mpg, though the moving average fluctuates over a range of approximately ±2 mpg.

FIGURE 14.20

Three-period moving average of MPG



Centered Moving Average (*CMA*)

Another moving average is the ***centered moving average*** (***CMA***). Formula 14.11 shows a *CMA* for $m = 3$ periods. The formula looks both forward *and* backward in time, to express the current "forecast" as the mean of the current observation *and* observations on either side of the current data.

(14.11) $$\hat{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3}$$ (centered moving average over m periods)

This is not really a forecast at all, but merely a way of smoothing the data. In Table 14.12, column seven shows the *CMA* for Andrew's MPG data. For example, for period 14 (blue-shaded cells) the *CMA* is

$$\hat{y}_t = \frac{24.198 + 24.505 + 19.210}{3} = 22.638$$

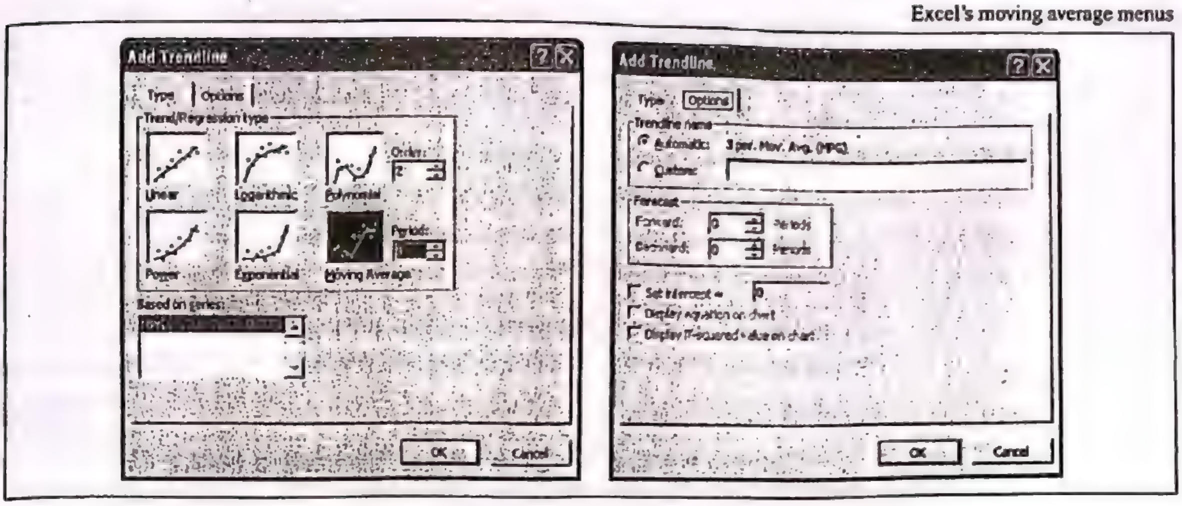
When n is odd ($m = 3, 5$, etc.) the *CMA* is easy to calculate. When m is even, the formula is more complex, because the mean of an even number of data points would lie *between* two data points and would not be correctly centered. Instead, we take a double moving average (yipe!) to get the resulting *CMA* centered properly. For example, for $m = 4$, we would average y_{t-1} through y_{t+1}, then average y_{t-1} through y_{t+2}, and finally average the two averages! You need not worry about this formula for now. It will be illustrated shortly in the context of seasonal data.

Using Excel for a TMA

Excel offers a *TMA* in its Add Trendline option when you click on a time-series line graph or bar chart. Its menus are displayed in Figure 14.21. The *TMA* is a conservative choice whenever you doubt that one of Excel's five other trend models (linear, logarithmic, polynomial, power, exponential) would be appropriate. However, Excel does *not* give you the option of making any forecasts with its moving average model.

FIGURE 14.21

Excel's moving average menus



SECTION EXERCISES

14.6 (a) Make an Excel line graph of the exchange rate data. Describe the pattern. (b) Click on the data and choose Add Trendline > Moving Average. Describe the effect of increasing m (e.g., $m = 2, 4, 6$, etc.). Include a copy of each graph with your answer. (c) Discuss how this moving average might help a currency speculator. DollarEuro

Daily Dollar/Euro Exchange Rate for First 3 Months of 2005 ($n = 64$ days)

Date	Rate	Date	Rate	Date	Rate	Date	Rate
3-Jan	1.3476	25-Jan	1.2954	16-Feb	1.2994	10-Mar	1.3409
4-Jan	1.3295	26-Jan	1.3081	17-Feb	1.3083	11-Mar	1.3465
5-Jan	1.3292	27-Jan	1.3032	18-Feb	1.3075	14-Mar	1.3346
6-Jan	1.3187	28-Jan	1.3033	21-Feb	1.3153	15-Mar	1.3315
7-Jan	1.3062	31-Jan	1.3049	22-Feb	1.3230	16-Mar	1.3423
10-Jan	1.3109	1-Feb	1.3017	23-Feb	1.3208	17-Mar	1.3373
11-Jan	1.3161	2-Feb	1.3015	24-Feb	1.3205	18-Mar	1.3311
12-Jan	1.3281	3-Feb	1.2959	25-Feb	1.3195	21-Mar	1.3165
13-Jan	1.3207	4-Feb	1.2927	28-Feb	1.3274	22-Mar	1.3210
14-Jan	1.3106	7-Feb	1.2773	1-Mar	1.3189	23-Mar	1.3005
17-Jan	1.3075	8-Feb	1.2783	2-Mar	1.3127	24-Mar	1.2957
18-Jan	1.3043	9-Feb	1.2797	3-Mar	1.3130	25-Mar	1.2954
19-Jan	1.3036	10-Feb	1.2882	4-Mar	1.3244	28-Mar	1.2877
20-Jan	1.2959	11-Feb	1.2864	7-Mar	1.3203	29-Mar	1.2913
21-Jan	1.3049	14-Feb	1.2981	8-Mar	1.3342	30-Mar	1.2944
24-Jan	1.3041	15-Feb	1.2986	9-Mar	1.3384	31-Mar	1.2969

Source: www.federalreserve.gov.

14.5 EXPONENTIAL SMOOTHING

Forecast Updating

The ***exponential smoothing*** model is a special kind of moving average. It is used for ongoing one-period-ahead forecasting for data that has up-and-down movements but no consistent trend. For example, a retail outlet may place orders for thousands of different stock-keeping units (SKUs) each week, so as to maintain its inventory of each item at the desired level (to avoid emergency calls to warehouses or suppliers). For such forecasts, many firms choose exponential smoothing, a simple forecasting model with only two inputs and one constant. The

updating formula for the forecasts is

(14.12) $$F_{t+1} = \alpha y_t + (1-\alpha)F_t \quad \text{(smoothing update)}$$

where

F_{t+1} = the forecast for the next period
α = the "smoothing constant" ($0 \le \alpha \le 1$)
y_t = the actual data value in period t
F_t = the previous forecast for period t

Smoothing Constant (α)

The next forecast F_{t+1} is a weighted average of y_t (the current data) and F_t (the previous forecast). The value of α, called the ***smoothing constant,*** is the weight given to the latest data. A small value of α would give low weight to the most recent observation and heavy weight $1-\alpha$ to the previous forecast (a "heavily smoothed" series). The larger the value of α, the more quickly the forecasts adapt to recent data. For example,

If $\alpha = .05$, then $F_{t+1} = .05y_t + .95F_t$ (heavy smoothing, slow adaptation)
If $\alpha = .20$, then $F_{t+1} = .20y_t + .80F_t$ (moderate smoothing, moderate adaptation)
If $\alpha = .50$, then $F_{t+1} = .50y_t + .50F_t$ (little smoothing, quick adaptation)

Choosing the Value of α

If $\alpha = 1$, there is no smoothing at all, and the forecast for next period is the same as the latest data point, which basically defeats the purpose of exponential smoothing. MINITAB uses $\alpha = .20$ (i.e., moderate smoothing) as its default, which is a fairly common choice of α. The fit of the forecasts to the data will change as you try different values of α. Most computer packages can, as an option, solve for the "best" α using a criterion such as minimum *SSE.*

Over time, earlier data values have less effect on the exponential smoothing forecasts than more recent y-values. To see this, we can replace F_t in equation 14.12 with the prior forecast F_{t-1}, and repeat this type of substitution indefinitely to obtain this result:

(14.13) $$F_{t+1} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \cdots$$

We see that the next forecast F_{t+1} depends on *all* the prior data (y_{t-1}, y_{t-2}, etc). As long as $\alpha < 1$, as we go farther into the past, each prior data value has less and less impact on the current forecast.

Initializing the Process

From equation 14.12, we see that F_{t+1} depends on F_t, which in turn depends on F_{t-1}, and so on, all the way back to F_1. But where do we get F_1 (the initial forecast)? There are many ways to initialize the forecasting process. For example, Excel simply sets the initial forecast equal to the first actual data value:

Method *A*

Set $F_1 = y_1$ (use the first data value)

This method has the advantage of simplicity, but if y_1 happens to be unusual, it could take a few iterations for the forecasts to stabilize. Another approach is to set the initial forecast equal to the average of the first several observed data values. For example, MINITAB uses the first six data values:

Method *B*

Set $F_1 = \dfrac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{n}$ (average of first 6 data values)

This method tends to iron out the effects of unusual y-values, but it consumes more data and is still vulnerable to unusual y-values.

Method *C*

Set F_1 = prediction from *backcasting* (backward extrapolation)

You may think of this method as fitting a trend to the data *in reverse time order* and extrapolating the trend to "predict" the initial value in the series. This method is common because it tends to generate a more appropriate initial forecast F_1. However, backcasting requires special software, so it will not be discussed here.

EXAMPLE

Weekly Sales Data

Table 14.13 shows weekly sales of deck sealer (a paint product sold in gallon containers) at a large do-it-yourself warehouse-style retailer. For exponential smoothing forecasts, the company uses $\alpha = .10$. Its choice of α is based on experience. Since α is fairly small, it will provide strong smoothing. The last two columns compare the two methods of initializing the forecasts. Unusually high sales in week 5 have a strong effect on method *B*'s starting point. At first, the difference in forecasts is striking, but over time the methods converge.

TABLE 14.13 Deck Sealer Sales: Exponential Smoothing ($n = 18$ weeks) DeckSealer

Week	*Sales in Gallons*	*Method A:* $F_1 = y_1$	*Method B:* F_1 = *Average (1st 6)*
1	106	106.000	127.833
2	110	106.000	125.650
3	108	106.400	124.085
4	97	106.560	122.477
5	210	105.604	119.929
6	136	116.044	128.936
7	128	118.039	129.642
8	134	119.035	129.478
9	107	120.532	129.930
10	123	119.179	127.637
11	139	119.561	127.174
12	140	121.505	128.356
13	144	123.354	129.521
14	94	125.419	130.969
15	108	122.277	127.272
16	168	120.849	125.344
17	179	125.564	129.610
18	120	130.908	134.549

Smoothed forecasts using $\alpha = .10$.

Using Method A:

$$F_2 = \alpha y_1 + (1-\alpha)F_1 = (.10)(106) + (.90)(106) = 106$$
$$F_3 = \alpha y_2 + (1-\alpha)F_2 = (.10)(110) + (.90)(106) = 106.4$$
$$F_4 = \alpha y_3 + (1-\alpha)F_3 = (.10)(108) + (.90)(106.4) = 106.56$$
$$\vdots$$
$$F_{19} = \alpha y_{18} + (1-\alpha)F_{18} = (.10)(120) + (.90)(130.908) = 129.82$$

Using Method B:

$$F_2 = \alpha y_1 + (1-\alpha)F_1 = (.10)(106) + (.90)(127.833) = 125.650$$
$$F_3 = \alpha y_2 + (1-\alpha)F_2 = (.10)(110) + (.90)(125.650) = 124.085$$
$$F_4 = \alpha y_3 + (1-\alpha)F_3 = (.10)(108) + (.90)(124.085) = 122.477$$
$$\vdots$$
$$F_{19} = \alpha y_{18} + (1-\alpha)F_{18} = (.10)(120) + (.90)(134.549) = 133.094$$

Despite their different starting points, the forecasts for period 19 do not differ greatly. Rounding to the next higher integer, for week 19, the firm would order 130 gallons (using method *A*) or 134 gallons (using method *B*). Figures 14.22 and 14.23 show the similarity in *patterns* of the forecasts, although the *level* of forecasts is always higher in method *B* because of its higher initial value. This demonstrates that the choice of starting values *does* affect the forecasts.

FIGURE 14.22

Using the first y-value

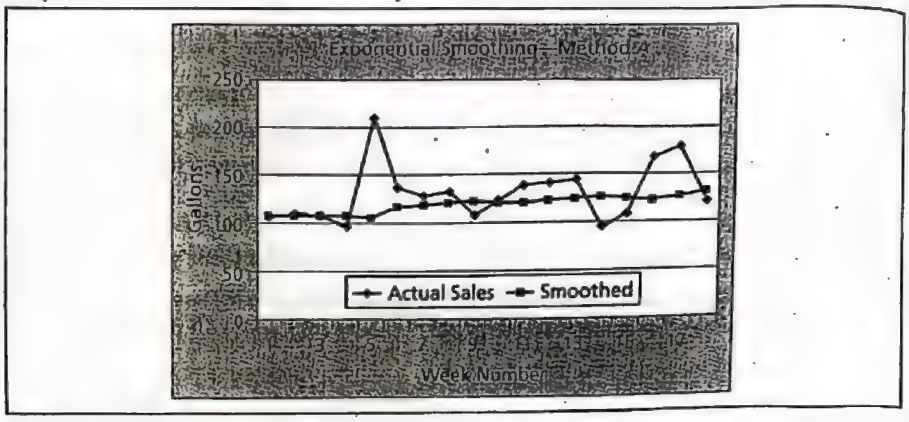
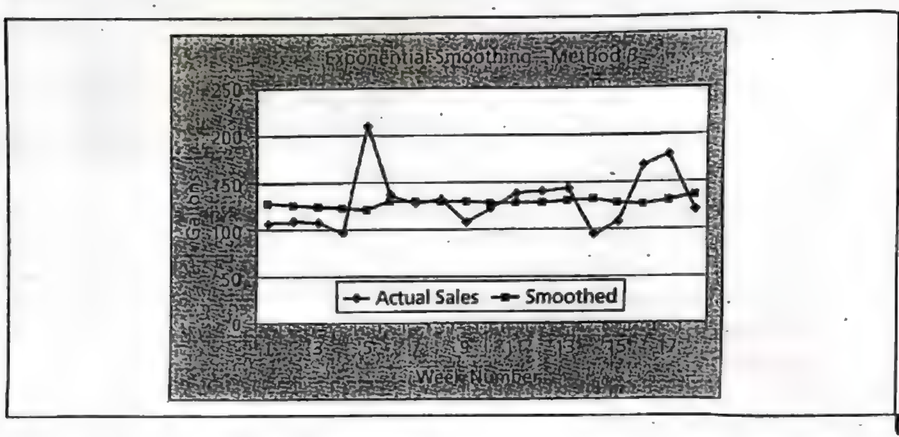


FIGURE 14.23

Averaging the first six y-values

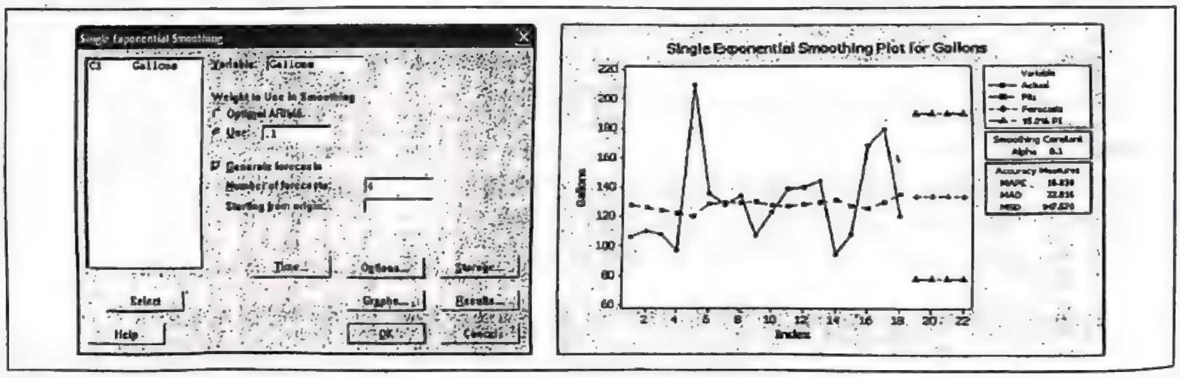


Using MINITAB

Figure 14.24 shows MINITAB's single exponential smoothing and 4 weeks' forecasts. After week 18, the exponential smoothing method cannot be updated with actual data, so the forecasts are constant. The wide 95 percent confidence intervals reflect the rather erratic past sales pattern.

FIGURE 14.24

MINITAB's exponential smoothing



Using Excel

Excel also has an exponential smoothing option. It is found in Data Analysis under the Tools menu. One difference to be noted is that Excel asks for a *damping factor,* which is equal to $1 - \alpha$. Excel uses method A to initialize the exponential smoothing forecasts. Figure 14.25 shows Excel's exponential smoothing dialogue box and a line chart of the actual values and forecast values. Notice that there are no forecast values beyond period 18 and that there are no confidence intervals as with MINITAB.

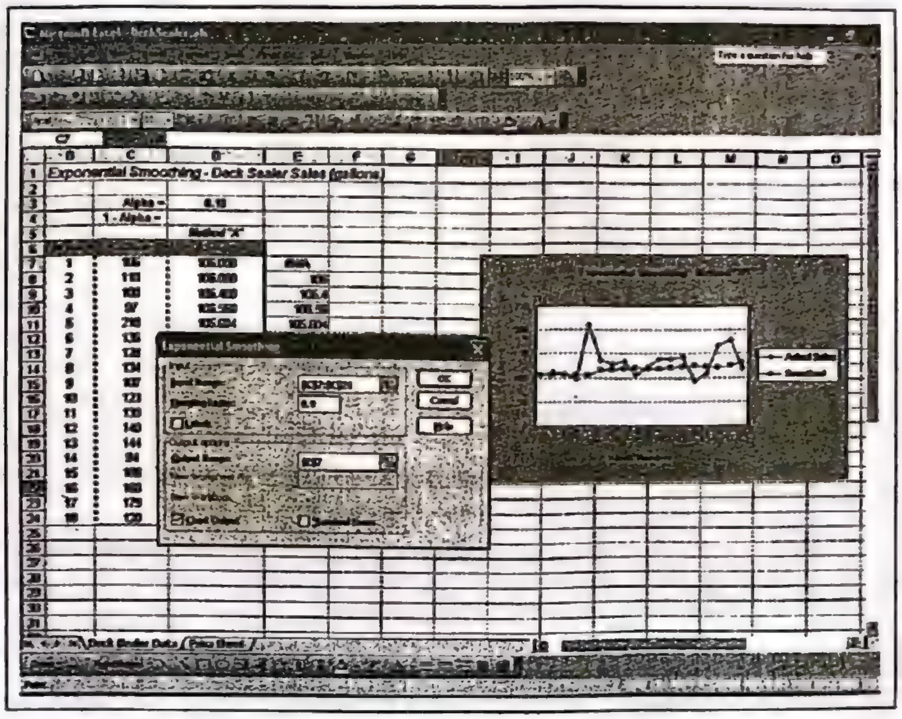


FIGURE 14.25
Excel's exponential smoothing

Smoothing with Trend and Seasonality

Single exponential smoothing is intended for *trendless* data. If your data have a trend, you can try *Holt's method* with *two* smoothing constants (one for *trend,* one for *level*). If you have both trend and seasonality, you can try *Winters's method* with *three* smoothing constants (one for *trend,* one for *level,* one for *seasonality*). These advanced methods are similar to single smoothing in that they use simple formulas to update the forecasts, and you may use them without special caution. *LearningStats* contains examples, explanations, and applications of these methods. Since these topics are usually reserved for a class in forecasting, they will not be explained here.

Mini Case 14.2

Exchange Rates

We have data for March 1 to March 30 and want to forecast 1 day ahead to March 31 by using exponential smoothing. We choose a smoothing constant value of $\alpha = .20$ and set the initial forecast F_1 to the average of the first six data values. Table 14.14 shows the actual data (y_t) and MINITAB's forecasts (F_t) for each date. The March 31 forecast is $F_{23} = \alpha y_{22} + (1-\alpha)F_{22} \doteq (.20)(1.2164) + (.80)(1.21395) = 1.2144$.

TABLE 14.14 Exchange Rate Canada/U.S. Dollar Canada

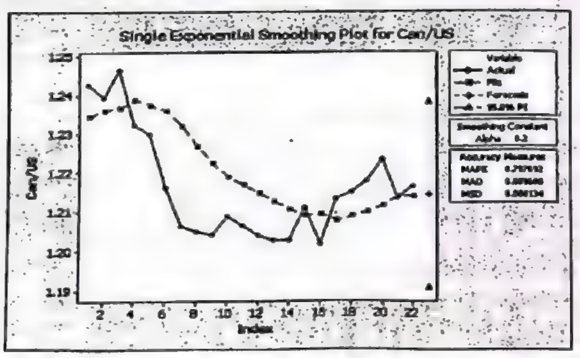
t	Date	y_t	F_t
1	1-Mar-05	1.2425	1.23450
2	2-Mar-05	1.2395	1.23610
3	3-Mar-05	1.2463	1.23678
4	4-Mar-05	1.2324	1.23868
5	7-Mar-05	1.2300	1.23743
6	8-Mar-05	1.2163	1.23594
7	9-Mar-05	1.2064	1.23201
8	10-Mar-05	1.2050	1.22689
9	11-Mar-05	1.2041	1.22251
10	14-Mar-05	1.2087	1.21883
11	15-Mar-05	1.2064	1.21680
12	16-Mar-05	1.2038	1.21472
13	17-Mar-05	1.2028	1.21254
14	18-Mar-05	1.2027	1.21059
15	21-Mar-05	1.2110	1.20901
16	22-Mar-05	1.2017	1.20941
17	23-Mar-05	1.2133	1.20787
18	24-Mar-05	1.2150	1.20895
19	25-Mar-05	1.2180	1.21016
20	28-Mar-05	1.2234	1.21173
21	29-Mar-05	1.2135	1.21406
22	30-Mar-05	1.2164	1.21395
23	31-Mar-05		1.21444

Source: www.federalreserve.gov.

Figure 14.26 shows MINITAB's plot of the data and forecasts. The forecasts adapt, but always with a lag. The actual exchange rate on March 31 was 1.2094, slightly lower than the forecast, but well within the 95 percent prediction limits.

FIGURE 14.26

MINITAB's exponential smoothing ($\alpha = .20$)



SECTION EXERCISES

14.7 (a) Make an Excel line graph of the following bond yield data. Describe the pattern. Is there a consistent trend? (b) Use exponential smoothing (MegaStat, MINITAB, or Excel) with $\alpha = .20$. Use both methods *A* and *B* to initialize the forecast (the default in both MegaStat and MINITAB). (c) Record the statistics of fit (MegaStat uses *MSE* and *MSD*, MINITAB uses *MSD* and *MAPE*). With Excel you will have to calculate these by creating cell formulas). (d) Do the smoothing again with $\alpha = .10$ and then with $\alpha = .30$, recording the statistics of fit. (e) Compare the statistics of fit for the three values of α. (f) Make a one-period forecast (i.e., $t = 53$) using each of the three α values. How did α affect your forecasts? **BondYield**

U.S. Treasury 10-Year Bond Yields at Week's End ($n = 52$ weeks)

Week	Yield	Week	Yield	Week	Yield	Week	Yield
4/2/04	3.95	7/2/04	4.63	10/1/04	4.10	12/31/04	4.29
4/9/04	4.21	7/9/04	4.49	10/8/04	4.20	1/7/05	4.28
4/16/04	4.36	7/16/04	4.47	10/15/04	4.08	1/14/05	4.25
4/23/04	4.43	7/23/04	4.46	10/22/04	4.03	1/21/05	4.19
4/30/04	4.49	7/30/04	4.56	10/29/04	4.05	1/28/05	4.19
5/7/04	4.62	8/6/04	4.41	11/5/04	4.12	2/4/05	4.14
5/14/04	4.81	8/13/04	4.28	11/12/04	4.22	2/11/05	4.06
5/21/04	4.74	8/20/04	4.23	11/19/04	4.17	2/18/05	4.16
5/28/04	4.68	8/27/04	4.25	11/26/04	4.20	2/25/05	4.28
6/4/04	4.74	9/3/04	4.19	12/3/04	4.35	3/4/05	4.37
6/11/04	4.80	9/10/04	4.21	12/10/04	4.19	3/11/05	4.45
6/18/04	4.75	9/17/04	4.14	12/17/04	4.16	3/18/05	4.51
6/25/04	4.69	9/24/04	4.04	12/24/04	4.21	3/25/05	4.59

Source: www.federalreserve.gov.

14.6 SEASONALITY

When and How to Deseasonalize

When the data periodicity is monthly or quarterly we should calculate a seasonal index and use it to ***deseasonalize*** the data (annual data have no seasonality). For a multiplicative model (the usual assumption) a seasonal index is a *ratio*. For example, if the seasonal index for July is 1.25, it means that July is 125 percent of the monthly average. If the seasonal index for January is 0.84, it means that January is 84 percent of the monthly average. If the seasonal index for October is 1.00, it means that October is an average month. The seasonal indexes must sum to 12 for monthly data or 4 for quarterly data. The following steps are used to deseasonalize data for time-series observations:

- Step 1 Calculate a centered moving average (*CMA*) for each month (quarter).
- Step 2 Divide each observed y_t value by the *CMA* to obtain seasonal ratios.
- Step 3 Average the seasonal ratios by month (quarter) to get raw seasonal indexes.
- Step 4 Adjust the raw seasonal indexes so they sum to 12 (monthly) or 4 (quarterly).
- Step 5 Divide each y_t by its seasonal index to get deseasonalized data.

In step 1, we lose 12 observations (monthly data) or 4 observations (quarterly data) because of the centering process. We will illustrate this technique for quarterly data.

Illustration of Calculations

Table 14.15 shows 6 years' data on quarterly revenue from sales of carpeting, tile, wood, and vinyl flooring by a floor-covering retailer. The data have an upward trend (see Figure 14.27), perhaps due to a boom in consumer spending on home improvement and new homes. There also appears to be seasonality, with lower sales in the third quarter (summer) and higher sales in the first quarter (winter).

TABLE 14.15 Sales of Floor Covering Materials ($ thousands) FloorSales

Quarter	2000	2001	2002	2003	2004	2005
1	259	306	379	369	515	626
2	236	300	262	373	373	535
3	164	189	242	255	339	397
4	222	275	296	374	519	488

The seasonal decomposition of this data is shown in Table 14.16 and Figure 14.27. Calculations are handled automatically by *MegaStat* so it's actually easy to perform the decomposition. Since the number of subperiods (quarters) is even ($m = 4$) each value of the *CMA* is the average of two averages. For example, the first *CMA* value 226.125 is the average of (259 + 236 + 164 + 222)/4 and (236 + 164 + 222 + 306)/4. Table 14.17 shows how the indexes are averaged. The *CMA* loses two quarters at the beginning and two quarters at the end, so each seasonal index is an average of only five quarters (instead of six). Each mean is then adjusted to force the sum to be 4.000, and these become the seasonal indexes. If we had monthly data, the indexes would be adjusted so that their sum would be 12.000.

TABLE 14.16
Calculation of Deseasonalized Sales ($n = 24$ quarters)
FloorSales

Obs	*Year*	*Quarter*	*Sales*	*CMA*	*Sales/CMA*	*Seasonal Index*	*Deseasonalized*
1	2000	1	259			1.252	206.9
2		2	236			1.021	231.1
3		3	164	*226.125*	0.725	0.740	221.7
4		4	222	240.000	0.925	0.987	224.9
5	2001	1	306	251.125	1.219	1.252	244.4
6		2	300	260.875	1.150	1.021	293.8
7		3	189	276.625	0.683	0.740	255.5
8		4	275	281.000	0.979	0.987	278.6
9	2002	1	379	282.875	1.340	1.252	302.7
10		2	262	292.125	0.897	1.021	256.6
11		3	242	293.500	0.825	0.740	327.2
12		4	296	306.125	0.967	0.987	299.8
13	2003	1	369	321.625	1.147	1.252	294.7
14		2	373	333.000	1.120	1.021	365.3
15		3	255	361.000	0.706	0.740	344.7
16		4	374	379.250	0.986	0.987	378.8
17	2004	1	515	389.750	1.321	1.252	411.3
18		2	373	418.375	0.892	1.021	365.3
19		3	339	450.375	0.753	0.740	458.3
20		4	519	484.500	1.071	0.987	525.7
21	2005	1	626	512.000	1.223	1.252	500.0
22		2	535	515.375	1.038	1.021	524.0
23		3	397			0.740	536.7
24		4	488			0.987	494.3

FIGURE 14.27
MegaStat's deseasonalized trend

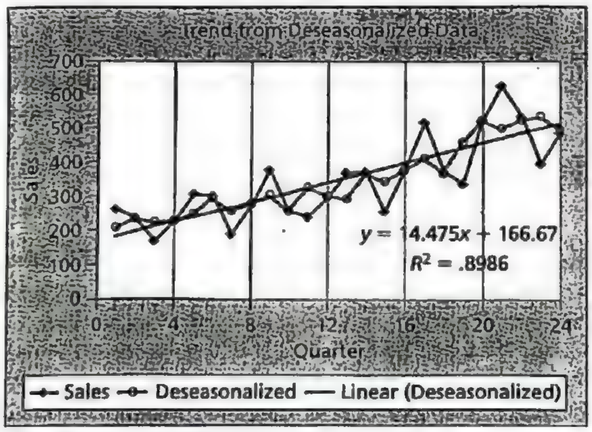


TABLE 14.17
Calculation of Seasonal Indexes
FloorSales

Quarter	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*Mean*	*Adjusted*
1		1.219	1.340	1.147	1.321	1.223	1.250	1.252
2		1.150	0.897	1.120	0.892	1.038	1.019	1.021
3	0.725	0.683	0.825	0.706	0.753		0.738	0.740
4	0.925	0.979	0.967	0.986	1.071		0.986	0.987
							3.993	4.000

Due to rounding, details may not yield the result shown.

After the data have been deseasonalized, the trend is fitted. Figure 14.27 shows the fitted trend from MegaStat, based on the deseasonalized data. The sharper peaks and valleys in the original time-series (Y) have been smoothed by removing the seasonality (S). Any remaining variation about the trend (T) is irregular (I) or "random noise."

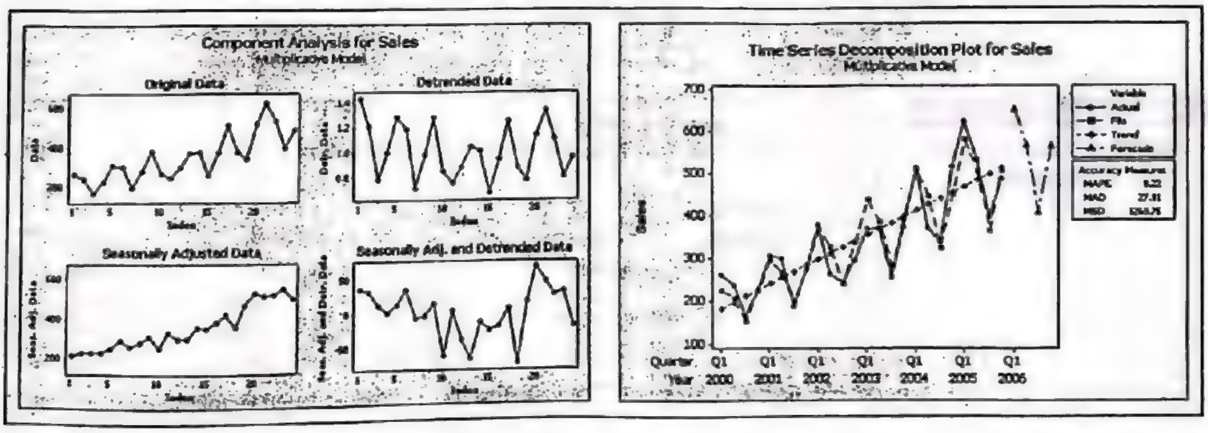
Using MINITAB to Deseasonalize

MINITAB performs its deseasonalization in a similar way, although it averages the seasonal factors using *medians* instead of *means,* so the results are not exactly the same as MegaStat's. For example, using the same floor covering sales data:

Quarter	*MegaStat's Seasonal Index*	*MINITAB's Seasonal Index*
1	1.252	1.234
2	1.021	1.047
3	0.740	0.732
4	0.987	0.987
Sum	4.000	4.000
Fitted trend	$y_t = 166.67 + 14.475t$	$y_t = 166.62 + 14.483t$

MINITAB offers nice graphical displays for decomposition, as well as forecasts, as shown in Figure 14.28. MINITAB also offers additive as well as multiplicative seasonality. In an additive model, the *CMA* is calculated in the same way, but the raw seasonals are *differences* (instead of ratios) and the seasonal indexes are forced to sum to *zero* (e.g., months with higher sales must exactly balance months with lower sales). Since most analysts prefer multiplicative models (assuming trended data) the additive model is not discussed in detail here.

FIGURE 14.28
MINITAB's graphs for floor covering sales



VS Chapter 19

Seasonal Forecasts Using Binary Predictors

Another way to address seasonality is to estimate a regression model using ***seasonal binaries*** as predictors. For quarterly data, for example, the data set would look as shown in Table 14.18. When we have four binaries (i.e., four quarters) we must exclude one binary to prevent perfect multicollinearity (see Chapter 13, Section 13.5). Arbitrarily, we exclude the fourth quarter binary *Qtr4* (it will be a portion of the intercept when $Qtr1 = 0$ and $Qtr2 = 0$ and $Qtr3 = 0$).

TABLE 14.18
Sales Data with Seasonal Binaries
FloorSales

Year	Quarter	Sales	Time	Qtr1	Qtr2	Qtr3
2000	1	259	1	1	0	0
	2	236	2	0	1	0
	3	164	3	0	0	1
	4	222	4	0	0	0
2001	1	306	5	1	0	0
	2	300	6	0	1	0
	3	189	7	0	0	1
	4	275	8	0	0	0
2002	1	379	9	1	0	0
	2	262	10	0	1	0
	3	242	11	0	0	1
	4	296	12	0	0	0
2003	1	369	13	1	0	0
	2	373	14	0	1	0
	3	255	15	0	0	1
	4	374	16	0	0	0
2004	1	515	17	1	0	0
	2	373	18	0	1	0
	3	339	19	0	0	1
	4	519	20	0	0	0
2005	1	626	21	1	0	0
	2	535	22	0	1	0
	3	397	23	0	0	1
	4	488	24	0	0	0

We assume a linear trend, and specify the regression model $Sales = f(Time, Qtr1, Qtr2, Qtr3)$. MINITAB's estimated regression is shown in Figure 14.29. This is an additive model of the form $Y = T + S + I$ (recall that we omit the cycle C in practice). The fitted equation is

$$Sales = 161 + 14.4\ Time + 89.8\ Qtr1 + 12.9\ Qtr2 - 83.6\ Qtr3$$

FIGURE 14.29
MINITAB's fitted regression for seasonal binaries

The regression equation is
Sales = 161 + 14.4 Time + 89.8 Qtr1 + 12.9 Qtr2 − 83.6 Qtr3

Predictor	Coef	SE Coef	T	P
Constant	161.21	24.33	6.62	0.000
Time	14.366	1.244	11.55	0.000
Qtr1	89.76	24.32	3.69	0.002
Qtr2	12.90	24.16	0.53	0.600
Qtr3	−83.63	24.07	−3.47	0.003

S = 41.6313   R-Sq = 90.0%   R-Sq(adj) = 87.9%

Time is a significant predictor ($p = .000$) indicating significant linear trend. Two of the binaries are significant: *Qtr1* ($p = .002$) and *Qtr3* ($p = .003$). The second quarter binary *Qtr2* ($p = .600$) is not significant. The model gives a good overall fit ($R^2 = .90$). The main virtue of the seasonal regression model is its versatility. We can plug in future values of *Time* and the seasonal binaries to create forecasts as far ahead as we wish. For example, the forecasts for 2006 are

$$\textit{Period 25: Sales} = 161 + 14.4(25) + 89.8(1) + 12.9(0) - 83.6(0) = 610.8$$
$$\textit{Period 26: Sales} = 161 + 14.4(26) + 89.8(0) + 12.9(1) - 83.6(0) = 548.3$$
$$\textit{Period 27: Sales} = 161 + 14.4(27) + 89.8(0) + 12.9(0) - 83.6(1) = 466.2$$
$$\textit{Period 28: Sales} = 161 + 14.4(28) + 89.8(0) + 12.9(0) - 83.6(0) = 564.2$$

SECTION EXERCISES

14.8 (a) Use MegaStat or MINITAB to deseasonalize the quarterly data on PepsiCo's revenues and fit a trend. Interpret the results. (b) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. (c) Use the regression equation to make a prediction for each quarter in 2005. (d) If you have access to *Standard & Poor's Stock Reports, 2006*, check your forecasts. How accurate were they? **PepsiCo**

PepsiCo Revenues ($ millions), 1998–2003

Quarter	1999	2000	2001	2002	2003	2004
1	5,114	4,191	5,330	5,101	5,530	6,131
2	4,982	4,928	6,713	6,178	6,538	7,070
3	4,591	4,909	6,906	6,376	6,830	7,257
4	5,680	6,410	7,986	7,457	8,073	8,803
Year	20,367	20,438	26,935	25,112	26,971	29,261

Source: *Standard & Poor's Stock Reports*, March 2005.

14.9 (a) Use MegaStat or MINITAB to deseasonalize the monthly Corvette sales data and fit a trend. Interpret the results. (b) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. (c) Use the regression equation to make a prediction for each month in 2004. (d) If you have access to *Ward's Automotive Yearbook, 2005* (67th edition), check your forecasts. How accurate were they? **Corvette**

U.S. Corvette Sales, 2000–2004 (number of cars sold)

Month	2000	2001	2002	2003
Jan	1,863	2,252	2,443	1,468
Feb	2,765	2,766	3,354	1,724
Mar	3,440	2,923	1,877	2,792
Apr	3,018	2,713	2,176	6,249
May	2,725	2,847	3,049	2,441
Jun	2,538	2,521	2,708	2,272
Jul	1,598	2,000	2,960	2,007
Aug	2,861	2,789	2,912	2,107
Sep	2,942	3,639	2,960	1,615
Oct	2,748	4,647	3,094	1,878
Nov	2,376	2,910	2,163	1,596
Dec	2,334	1,648	2,859	1,825
Total	31,208	33,655	32,555	27,974

Source: *Ward's Automotive Yearbook, 2001–2004.*

Mini Case 14.3

Beer Shipments Beer

Table 14.19 shows U.S. beer shipments by month for 1995–2000. To analyze trend and seasonality, we create a regression data set with linear trend (*Time* = 1, 2, . . . , 72) and 11 seasonal binaries (e.g., *Jan* = 1 if it's January, 0 otherwise). The December binary is omitted to prevent perfect multicollinearity.

TABLE 14.19 U.S. Beer Shipments, 1995–2000 (thousands of gross) Beer

Month	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*
Jan	8,635	8,606	9,161	9,574	9,673	9,827
Feb	8,179	8,577	8,774	9,098	9,757	9,907
Mar	9,820	9,830	10,198	10,263	11,647	11,067
Apr	8,735	10,188	10,499	10,160	10,834	10,599
May	10,332	11,289	11,022	10,871	11,337	11,710
Jun	10,336	9,933	11,034	11,812	12,034	11,799
Jul	9,864	11,233	11,169	11,679	10,958	11,279
Aug	10,182	10,258	10,373	10,692	10,717	11,537
Sep	9,422	9,249	10,143	10,165	10,406	10,412
Oct	9,671	9,913	9,822	9,917	9,755	10,512
Nov	8,469	8,742	8,895	9,528	10,204	9,874
Dec	7,385	8,077	9,091	8,963	9,373	9,007
Total	111,030	115,895	120,181	122,722	126,695	127,530

Note: One gross equals 144 bottles.
Source: An independent project by statistics student Mai Lee using data from The U.S. Dept. of Commerce.

The regression results, shown in Figure 14.30, indicate a good fit ($R^2 = .892$), significant upward trend ($p = .000$ for *Time*), and significant seasonal binaries (all have very small p-values). The coefficients of the monthly binaries indicate high beer sales in May, June, and July, presumably because people drink more beer in hot weather.

FIGURE 14.30

MINITAB's fitted regression for seasonal binaries

The regression equation is
Beer = 7671 + 23.3 Time + 853 Jan + 632 Feb + 2031 Mar + 1706 Apr
+ 2607 May + 2648 Jun + 2498 Jul + 2070 Aug + 1387 Sep
+ 1329 Oct + 659 Nov

Predictor	Coef	SE Coef	T	P
Constant	7670.7	172.4	44.50	0.000
Time	23.302	2.090	11.15	0.000
Jan	853.0	211.1	4.04	0.000
Feb	632.4	210.8	3.00	0.004
Mar	2031.2	210.6	9.64	0.000
Apr	1706.2	210.5	8.11	0.000
May	2607.3	210.3	12.40	0.000
Jun	2648.5	210.2	12.60	0.000
Jul	2497.5	210.1	11.89	0.000
Aug	2070.4	210.0	9.86	0.000
Sep	1386.7	209.9	6.61	0.000
Oct	1328.9	209.8	6.33	0.000
Nov	659.3	209.8	3.14	0.003

S = 363.379   R-Sq = 89.2%   R-Sq (adj) = 87.0%

14.7 FORECASTING: FINAL THOUGHTS

Role of Forecasting

In many ways, forecasting resembles planning. *Forecasting* is an analytical way to describe a "what-if" future that might confront the organization. *Planning* is the organization's attempt to determine a set of actions it will take under each foreseeable contingency. Forecasts help decision makers become aware of trends or patterns that will require a response. Actions taken by the decision makers may actually head off the contingency envisioned in the forecast. Thus, forecasts tend to be self-defeating because they trigger homeostatic organizational responses.

Behavioral Aspects of Forecasting

Forecasts can facilitate organizational communication. The forecast (or even just a nicely prepared time-series chart) lets everyone examine the same facts concurrently, and perhaps argue with the data or the assumptions that underlie the forecast or its relevance to the organization. A quantitative forecast helps *make assumptions explicit*. Those who prepare the forecast must explain and defend their assumptions, while others must challenge them. In the process, everyone gains understanding of the data, the underlying realities, and the imperfections in the data. Forecasts *focus the dialogue* and can make it more productive.

Of course, this assumes a certain maturity among the individuals around the table. Strong leaders (or possibly meeting facilitators) can play a role in guiding the discourse to produce a positive result. The danger is that people may try to find scapegoats (yes, they do tend to blame the forecaster), deny facts, or avoid responsibility for tough decisions. But one premise of this book is that statistics, when done well, can strengthen any dialogue and lead to better decisions.

Forecasts Are Always Wrong

We discussed several measures to use to determine if a forecast model fits the time series. Successful forecasters understand that a forecast is never precise. There is always some error, but we can *use* the error measures to track forecast error. Many companies use several different forecasting models and rely on the model that has had the least error over some time period. We have described simple models in this chapter. You may take a class specifically focusing on forecasting in which you will learn about other time-series models including AR (autoregressive) models. AR models take advantage of the dependency that might exist between values in the time series, and belong to a class of models called ARIMA (autoregressive integrated moving average) models.

To ensure good forecast outcomes

- Maintain up-to-date databases of *relevant* data.
- Allow sufficient lead time to analyze the data.
- State several alternative forecasts or scenarios.
- Track forecast errors over time.
- State your assumptions and qualifications.
- Bear in mind the purpose of the forecasts.
- Consider the time horizon for the decision.
- Don't underestimate the power of a good graph.

There is always a role for "judgment" forecasts when time is short, patterns are unclear, or you have erratic or low-quality data. Watch out for unbelievable forecasts—they may be telling you that something is wrong somewhere. Don't try to dazzle people with equations that are not helpful. Consider ignoring the earlier part of the time-series if the series is long. And remember the principle of Occam's Razor.

Principle of Occam's Razor

Given two *sufficient* explanations, we prefer the simpler one.

William of Occam (1285–1347)

Chapter Summary

A **time series** is assumed to have four components. For most business data, **trend** is the general pattern of change over all years observed while **cycle** is a repetitive pattern of change around the trend over several years and **seasonality** is a repetitive pattern within a year. The **irregular** component is a random disturbance that follows no pattern. The **additive model** is adequate in the short run because the four components' magnitude does not change much, but for observations over longer periods of time, the **multiplicative model** is preferred. Common trend models include **linear** (constant slope and no turning point), **quadratic** (one turning point), and **exponential** (constant percent growth or decline). Higher polynomial models are untrustworthy and liable to give strange forecasts, though any trend model is less reliable the farther out it is projected. In forecasting, forecasters use fit measures besides R^2, such as mean absolute percent error (*MAPE*), mean absolute deviation (*MAD*), and mean squared deviation (*MSD*). For trendless or erratic data, we use a **moving average** over *m* periods or **exponential smoothing**. Forecasts adapt rapidly to changing data when the **smoothing constant** α is large (near 1) and conversely for a small α (near 0). For monthly or quarterly data, a **seasonal adjustment** is required before extracting the trend. Alternatively, regression with **seasonal binaries** can be used to capture seasonality and make forecasts.

Key Terms

Commonly Used Formulas

Additive time-series model: $Y = T + C + S + I$

Multiplicative time-series model: $Y = T \times C \times S \times I$

Linear trend model: $y_t = a + bt$

Exponential trend model: $y_t = ae^{bt}$

Quadratic trend model: $y_t = a + bt + ct^2$

Coefficient of determination: $R^2 = 1 - \dfrac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$

Mean absolute percent error: $MAPE = \dfrac{100}{n}\sum_{t=1}^{n}\dfrac{|y_t - \hat{y}_t|}{y_t}$

Mean absolute deviation: $MAD = \dfrac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$

Mean squared deviation: $MSD = \dfrac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$

Standard error: $SE = \sqrt{\sum_{t=1}^{n}\dfrac{(y_t - \hat{y}_t)^2}{n-2}}$

Forecast updating equation for exponential smoothing: $F_{t+1} = \alpha y_t + (1-\alpha)F_t$

Chapter Review

Note: Questions marked with an asterisk * refer to optional material.

1. Explain the difference between (a) stocks and flows; (b) cross-sectional and time-series data; (c) additive and multiplicative models.
2. (a) What is periodicity? (b) Give original examples of data with different periodicity.
3. (a) What are the distinguishing features of each component of a time series (trend, cycle, seasonal, irregular)? (b) Why is cycle usually ignored in time-series modeling?
4. Name four criteria for assessing a trend forecast.
5. Name two advantages and two disadvantages of each of the common trend models (linear, exponential, quadratic).
6. When would the exponential trend model be preferred to a linear trend model?
7. Explain how to obtain the compound percent growth rate from a fitted exponential model.
8. (a) When might a quadratic model be useful? (b) What precautions must be taken when forecasting with a quadratic model? (c) Why are higher-order polynomial models dangerous?
9. Name five measures of fit for a trend, and state their advantages and disadvantages.
10. (a) When do we use a moving average? (b) Name two types of moving averages. (c) When is a centered moving average harder to calculate?
11. (a) When is exponential smoothing most useful? (b) Interpret the smoothing constant α. What is its range? (c) What does a small α say about the degree of smoothing? A large α?
12. (a) Explain two ways to initialize the forecasts in an exponential smoothing process. (b) Name an advantage and a disadvantage of each method.
13. (a) Why is seasonality irrelevant for annual data? (b) List the steps in deseasonalizing a monthly time series. (c) What is the sum of a monthly seasonal index? A quarterly index?
14. (a) How can forecasting improve communication within an organization? (b) List five tips for ensuring effective forecasting outcomes.

*15. (a) Explain how seasonal binaries can be used to model seasonal data. (b) What is the advantage of using seasonal binaries?

*16. Explain the equivalency between the two forms of an exponential trend model.

CHAPTER EXERCISES

Instructions: For each exercise, use Excel, MegaStat, or MINITAB to make an attractive, well-labeled time-series line chart. Adjust the Y-axis scale if necessary to show more detail (since Excel usually starts the scale at zero). If a fitted trend is called for, use Excel's option to display the equation and R^2 statistic (or *MAPE, MAD,* and *MSD* in MINITAB). Include printed copies of all relevant graphs with your answers to each exercise.

14.10 (a) Choose *one* time-series describing Spirit Airlines and make a line chart. (b) Describe the trend (if any) and discuss possible causes. (c) Fit both a linear and an exponential trend to the data. (d) Which model is preferred? Why? (e) Make a forecast for 2003, using a trend model of your choice (or a judgment forecast). Spirit

Growth of Spirit Airlines, 1998–2002

Year	*Revenue ($ mil)*	*Aircraft*	*FT Employees*
1998	131	14	860
1999	227	20	1,440
2000	311	24	1,729
2001	354	27	2,094
2002	403	28	2,345

Source: *Detroit Free Press*, August 21, 2003, p. F1.

14.11 (a) Plot both Swiss watch time series on the same graph. (b) Describe the trend (if any) and discuss possible causes. (c) Fit an exponential trend to each time series. (d) Interpret each fitted trend carefully. What conclusion do you draw? (e) Make forecasts for 2003, using the linear trend model. Do you feel confident in your forecasts? Explain. **Swiss**

Swiss Watch Exports (thousands of units), 1998–2003

Year	*Mechanical*	*Electronic*
1998	2,558	29,678
1999	2,526	28,766
2000	2,549	27,313
2001	2,580	23,811
2002	2,722	24,107
2003	2,718	21,864

Source: Fédération de L'Industrie Horlogère Suisse, Swiss Watch Exports, www.fhs.ch.

14.12 (a) Plot the total minutes of TV viewing time per household. (b) Describe the trend (if any) and discuss possible causes. (c) Fit a linear trend to the data. (d) Would this model give reasonable forecasts? Would another trend model be better? Explain. (e) Make a forecast for 2005. Check the forecast if you have access to the Web. Show the forecast calculations. (f) Would this data ever approach an asymptote? Explain. *Note:* Time is in 5-year increments, so use $t = 12$ for the 2005 forecast. **Television**

Average Daily TV Viewing Time Per U.S. Household

Year	*Hours*	*Min*	*Total Min*
1950	4	35	275
1955	4	51	291
1960	5	6	306
1965	5	29	329
1970	5	56	356
1975	6	7	367
1980	6	36	396
1985	7	10	430
1990	6	53	413
1995	7	17	437
2000	7	35	455

Source: As published by the TVB based on Nielsen Media Research data. Used with permission.

14.13 (a) Plot the voter participation rate. (b) Describe the trend (if any) and discuss possible causes. (c) Fit both a linear and an exponential trend to the data. (d) Which model is preferred? Why? (e) Make a forecast for 2004, using a trend model of your choice (or a judgment forecast). (f) Check the Web for the actual 2004 voter participation rate. How close was your forecast? *Note:* Time is in 4-year increments, so use $t = 19$ for the 2004 forecast. **Voters**

U.S. Presidential Election Voter Participation, 1932–2000

Year	Voting Age Population	Voted for President	% Voting Pres
1932	75,768	39,758	52.5
1936	80,174	45,654	56.9
1940	84,728	49,900	58.9
1944	85,654	47,977	56.0
1948	95,573	48,794	51.1
1952	99,929	61,551	61.6
1956	104,515	62,027	59.3
1960	109,672	68,838	62.8
1964	114,090	70,645	61.9
1968	120,285	73,212	60.9
1972	140,777	77,719	55.2
1976	152,308	81,556	53.5
1980	163,945	86,515	52.8
1984	173,995	92,653	53.3
1988	181,956	91,595	50.3
1992	189,524	104,425	55.1
1996	196,928	96,278	49.0
2000	207,884	105,397	50.7

Source: *Statistical Abstract of the United States, 2001*, www.census.gov.

14.14 (a) Plot the market-share data. (b) Describe the trend (if any) and discuss possible causes. (c) Fit three trends (linear, exponential, quadratic). (d) Which trend model is best, and why? If none is satisfactory, explain. (e) Make a forecast for 2004 by using a trend model of your choice or a judgment forecast. **Trucks**

Asian and European Share of U.S. Light Truck Sales, 1990–2003

Year	Percent	Year	Percent
1990	16.4	1997	15.4
1991	17.1	1998	16.2
1992	14.3	1999	18.4
1993	13.7	2000	21.2
1994	14.2	2001	23.1
1995	13.6	2002	23.9
1996	13.6	2003	26.6

Source: *Detroit Free Press*, November 19, 2003, p. 1A.

14.15 (a) Choose *one* category of consumer credit and plot it. (b) Describe the trend (if any) and discuss possible causes. (c) Fit a trend model of your choice. (d) Make a forecast for 2004, using a trend model of your choice. *Note:* Revolving credit is mostly credit card and home equity loans, while nonrevolving credit is for a specific purchase such as a car. **Consumer**

Consumer Credit Outstanding, 1994–2003 ($ billions)

Year	Total	Revolving	Nonrevolving
1995	1,141	443	698
1996	1,242	499	743
1997	1,305	522	783
1998	1,400	563	837
1999	1,513	590	922
2000	1,686	659	1,027
2001	1,822	704	1,118
2002	1,903	717	1,186
2003	2,002	745	1,257

Source: *Statistical Abstract of the United States, 2004*.

14.16 (a) Plot the data on U.S. general aviation shipments. (b) Describe the pattern and discuss possible causes. (c) Would a fitted trend be helpful? Explain. (d) Make a similar graph for 1992–2003 only. Would a fitted trend be helpful in making a prediction for 2004? (e) Fit a trend model of your choice to the 1992–2003 data. (f) Make a forecast for 2004, using either the fitted trend model or a judgment forecast. Why is it best to ignore earlier years in this data set? **Airplanes**

U.S. Manufactured General Aviation Shipments, 1966–2003

Year	*Planes*	*Year*	*Planes*	*Year*	*Planes*	*Year*	*Planes*
1966	15,587	1976	15,451	1986	1,495	1996	1,053
1967	13,484	1977	16,904	1987	1,085	1997	1,482
1968	13,556	1978	17,811	1988	1,143	1998	2,115
1969	12,407	1979	17,048	1989	1,535	1999	2,421
1970	7,277	1980	11,877	1990	1,134	2000	2,714
1971	7,346	1981	9,457	1991	1,021	2001	2,538
1972	9,774	1982	4,266	1992	856	2002	2,169
1973	13,646	1983	2,691	1993	870	2003	2,090
1974	14,166	1984	2,431	1994	881		
1975	14,056	1985	2,029	1995	1,028		

Source: U.S. Manufactured General Aviation Shipments, *Statistical Databook 2003*, General Aviation Manufacturers Association, used with permission.

14.17 (a) Choose *one* beverage category and plot the data. (b) Describe the trend (if any) and discuss possible causes. (c) Would a fitted trend be helpful? Explain. (d) Fit several trend models. Which is best, and why? If none is satisfactory, explain. (e) Make a forecast for 2005, using a trend model of your choice or a judgment forecast. Discuss. *Note:* Time increments are 5 years, so use $t = 6$ for your 2005 forecast. **Beverages**

U.S. Per Capita Annual Consumption of Selected Beverages (gallons)

Beverage	*1980*	*1985*	*1990*	*1995*	*2000*
Milk	27.6	26.7	25.7	23.9	22.5
Whole	17.0	14.3	10.5	8.6	8.1
Reduced-fat	10.5	12.3	15.2	15.3	14.4
Carbonated soft drinks	35.1	35.7	46.2	47.4	49.3
Diet	5.1	7.1	10.7	10.9	11.6
Regular	29.9	28.7	35.6	36.5	37.7
Fruit juices	7.4	7.8	7.8	8.3	8.7
Alcoholic	28.3	28.0	27.5	24.7	24.9
Beer	24.3	23.8	23.9	21.8	21.7
Wine	2.1	2.4	2.0	1.7	2.0
Distilled spirits	2.0	1.8	1.5	1.2	1.3

Source: *Statistical Abstract of the United States, 2003.*

14.18 (a) Plot *either* receipts and outlays *or* federal debt and GDP (plot both time series on the same graph). (b) Describe the trend (if any) and discuss possible causes. (c) Fit an exponential trend to each. (d) Interpret each fitted trend equation, explaining its implications. (e) To whom is this issue relevant? **FedBudget**

U.S. Federal Finances, 1990–2004 ($ billions current)

Year	Receipts	Outlays	Federal Debt	GDP
1990	1,032	1,253	3,206	5,803
1991	1,055	1,324	3,598	5,996
1992	1,091	1,382	4,002	6,338
1993	1,154	1,410	4,351	6,657
1994	1,259	1,462	4,643	7,072
1995	1,352	1,516	4,921	7,398
1996	1,453	1,561	5,182	7,817
1997	1,579	1,601	5,369	8,304
1998	1,722	1,653	5,478	8,747
1999	1,828	1,702	5,606	9,268
2000	2,025	1,789	5,629	9,817
2001	1,991	1,863	5,770	10,128
2002	1,853	2,011	6,198	10,487
2003	1,782	2,160	6,760	11,004
2004	1,880	2,292	7,355	11,728

Source: *Economic Report of the President, 2004.*

14.19 (a) Plot both men's and women's winning times on the same graph. (b) Fit a linear trend model to each series. From the fitted trends, will the times eventually converge? *Hint:* Ask Excel for forecasts (e.g., 20 years ahead). (c) Make a copy of your graph, and click each fitted trend and change it to a moving average trend type. (d) Would a moving average be a reasonable approach to modeling these data sets? *Note:* The data file **Boston** has the data converted to decimal minutes.

Boston Marathon Champions, 1980–2005

	Men		Women	
Year	Name of Winner	Time	Name of Winner	Time
1980	Bill Rodgers	2:12:11	Jacqueline Gareau	2:34:28
1981	Toshihiko Seko	2:09:26	Allison Roe	2:26:46
1982	Alberto Salazar	2:08:52	Charlotte Teske	2:29:33
1983	Greg Meyer	2:09:00	Joan Benoit	2:22:43
1984	Geoff Smith	2:10:34	Lorraine Moller	2:29:30
1985	Geoff Smith	2:14:05	Lisa Larsen Weidenbach	2:34:10
1986	Robert de Castella	2:07:51	Ingrid Kristiansen	2:24:55
1987	Toshihiko Seko	2:11:50	Rosa Mota	2:25:21
1988	Ibrahim Hussein	2:08:43	Rosa Mota	2:24:30
1989	Abebe Mekonnen	2:09:06	Ingrid Kristiansen	2:24:33
1990	Gelindo Bordin	2:08:19	Rosa Mota	2:25:24
1991	Ibrahim Hussein	2:11:06	Wanda Panfil	2:24:18
1992	Ibrahim Hussein	2:08:14	Olga Markova	2:23:43
1993	Cosmas Ndeti	2:09:33	Olga Markova	2:25:27
1994	Cosmas Ndeti	2:07:15	Uta Pippig	2:21:45
1995	Cosmas Ndeti	2:09:22	Uta Pippig	2:25:11
1996	Moses Tanui	2:09:15	Uta Pippig	2:27:12
1997	Lameck Aguta	2:10:34	Fatuma Roba	2:26:23
1998	Moses Tanui	2:07:34	Fatuma Roba	2:23:21
1999	Joseph Chebet	2:09:52	Fatuma Roba	2:23:25
2000	Elijah Lagat	2:09:47	Catherine Ndereba	2:26:11
2001	Lee Bong-Ju	2:09:43	Catherine Ndereba	2:23:53
2002	Rodgers Rop	2:09:02	Margaret Okayo	2:20:43
2003	Robert Kipkoech Cheruiyot	2:10:11	Svetlana Zakharova	2:25:20
2004	Timothy Cherigat	2:10:37	Catherine Ndereba	2:24:27
2005	Hailu Negussie	2:11:45	Catherine Ndereba	2:25:13

Source: www.boston.com/marathon/history.

14.20 (a) Choose *either* commercial banks *or* savings institutions. On the same graph, plot both the main and branch data. (b) Fit a linear trend to each. (c) Interpret each fitted linear trend equation, explaining its implications for bank customers. (d) Make a copy of your graph, click each trend line, and change the trend type to exponential. (e) Interpret each fitted exponential trend. (f) Which is preferable, the linear or exponential trend model? FDIC

Number of FDIC-Insured Financial Institutions, 1995–2003

	Commercial Banks			Savings Institutions		
Year	*Banks*	*Main*	*Branches*	*Institutions*	*Main*	*Branches*
1995	65,888	9,971	55,917	15,462	2,030	13,432
1996	66,810	9,553	57,258	15,767	1,926	13,841
1997	68,810	9,165	59,645	14,831	1,780	13,051
1998	70,052	8,793	61,259	14,535	1,690	12,845
1999	71,534	8,597	62,937	14,506	1,642	12,864
2000	71,911	8,331	63,580	14,041	1,589	12,452
2001	72,458	8,095	64,363	14,048	1,534	12,514
2002	74,072	7,887	66,185	13,765	1,467	12,298
2003	75,159	7,769	67,390	13,937	1,413	12,524

Source: *Statistical Abstract of the United States, 2004.*

14.21 (a) Plot the data on fractional ownership of aircraft (i.e., shared ownership of available flight time). (b) Describe the trend (if any) and discuss possible causes. *Hint:* If you do not know what fractional ownership of aircraft is, use Google. (c) Fit the exponential trend to the data. Would this model give reasonable forecasts? Explain. (d) Make a forecast for 2003, using a trend model of your choice, or a judgment forecast. Fractional

Fractional Shares of Aircraft Ownership, 1986–2002

Year	*Shares*
1986	3
1987	5
1988	26
1989	51
1990	57
1991	71
1992	84
1993	110
1994	158
1995	285
1996	548
1997	957
1998	1551
1999	2607
2000	3834
2001	4871
2002	5827

Source: Chris Martin, David Jones, and Pinar Keskinocak, "Optimizing On-Demand Aircraft Schedules for Fractional Aircraft Operators," *Interfaces* 33, no. 5 (Sept.–Oct. 2003), p. 23.

14.22 (a) Plot all four time series on fuel efficiency on the same graph. (b) Fit a linear trend to each. (c) Interpret each fitted trend equation, explaining its implications. (d) To whom is this issue relevant? *Note:* If you think your graph is too cluttered, break it into two graphs (existing vehicles, new vehicles) with two time series on each graph. FuelMPG

Average Fuel Efficiency of U.S. Passenger Cars and Light Trucks (miles per gallon)

	Existing Vehicles		New Vehicles	
Year	*Passenger Car*	*Other Vehicles*	*Car*	*Light Truck*
1990	20.3	16.1	28.0	20.8
1991	21.2	17.0	28.4	21.3
1992	21.0	17.3	27.9	20.8
1993	20.6	17.4	28.4	21.0
1994	20.8	17.3	28.3	20.8
1995	21.1	17.3	28.6	20.5
1996	21.2	17.2	28.5	20.8
1997	21.5	17.2	28.7	20.6
1998	21.6	17.2	28.8	21.1
1999	21.4	17.0	28.3	20.9
2000	21.9	17.4	28.5	21.3
2001	22.1	17.6	28.6	20.9

Source: U.S. Dept. of Transportation, www.bts.gov.

14.23 (a) Plot the data on law enforcement officers killed. (b) Describe the trend (if any) and discuss possible causes. (c) Would a fitted trend be helpful? Explain. (c) Make a forecast for 2002 using any method you like (including judgment). **LawOfficers**

U.S. Law Enforcement Officers Killed, 1994–2002

Year	*Officers Killed*
1994	141
1995	133
1996	113
1997	133
1998	142
1999	107
2000	134
2001	218
2002	132

Source: *Statistical Abstract of the United States, 2004.*

14.24 (a) Plot the data on lightning deaths. (b) Describe the trend (if any) and discuss possible causes. (c) Fit an exponential trend to the data. Interpret the fitted equation. (d) Make a forecast for 2005, using a trend model of your choice (or a judgment forecast). Explain the basis for your forecast. *Note:* Time is in 5-year increments, so use $t = 14$ for your 2005 forecast. **Lightning**

U.S. Lightning Deaths, 1940–2000

Year	*Deaths*
1940	340
1945	268
1950	219
1955	181
1960	129
1965	149
1970	122
1975	91
1980	74
1985	74
1990	74
1995	85
2000	51

Source: *Statistical Abstract of the United States, 2003*, and *U.S. News & World Report* 108, no. 22 (June 4, 1990), p. 78.

14.25 (a) Plot the data on full-time mathematics graduate students. (b) Would a fitted trend be helpful? Explain. (c) Make a forecast for 2003, using a trend model of your choice (or a judgment forecast). MathGrads

Full-Time Mathematics Graduate Students, 1993–2002

Year	Total
1993	10,525
1994	10,185
1995	9,761
1996	9,476
1997	9,003
1998	8,791
1999	8,838
2000	9,637
2001	9,361
2002	9,972

Source: American Mathematical Association.

14.26 (a) Plot both men's and women's winning times on the same graph. (b) Fit a linear trend model to each series (men, women). (c) Use Excel's option to forecast each trend graphically to 2040 (i.e., to period $t = 27$ periods, since observations are in 4-year increments). From these projections, does it appear that the times will eventually converge? *(d) Set the fitted trends equal, solve for x (the time period when the trends will cross), and convert x to a year. Is the result plausible? Explain. (e) Use the Web to check your 2004 forecasts. Olympic

Summer Olympics 100-Meter Winning Times

Year	Men's 100-Meter Winner	Seconds	Women's 100-Meter Winner	Seconds
1928	Percy Williams, Canada	10.80	Elizabeth Robinson, United States	12.20
1932	Eddie Tolan, United States	10.30	Stella Walsh, Poland	11.90
1936	Jesse Owens, United States	10.30	Helen Stephens, United States	11.50
1948	Harrison Dillard, United States	10.30	Fanny Blankers-Koen, Netherlands	11.90
1952	Lindy Remigino, United States	10.40	Marjorie Jackson, United States	11.50
1956	Bobby Morrow, United States	10.50	Betty Cuthbert, Australia	11.50
1960	Armin Hary, West Germany	10.20	Wilma Rudolph, United States	11.00
1964	Bob Hayes, United States	10.00	Wyomia Tyus, United States	11.40
1968	Jim Hines, United States	9.95	Wyomia Tyus, United States	11.00
1972	Valery Borzov, USSR	10.14	Renate Stecher, East Germany	11.07
1976	Hasely Crawford, Trinidad	10.06	Annegret Richter, West Germany	11.08
1980	Allan Wells, Great Britain	10.25	Lyudmila Kondratyeva, USSR	11.06
1984	Carl Lewis, United States	9.99	Evelyn Ashford, United States	10.97
1988	Carl Lewis, United States	9.92	Florence Griffith-Joyner, United States	10.54
1992	Linford Christie, Great Britain	9.96	Gail Devers, United States	10.82
1996	Donovan Bailey, Canada	9.84	Gail Devers, United States	10.94
2000	Maurice Greene, United States	9.87	Marion Jones, United States	10.75

Source: Summer Olympics 100-meter times, *The World Almanac, 2002*, pp. 900–904.

14.27 (a) Choose *one* time series on U.S. petroleum use, and plot it on a graph. (b) Describe the trend (if any) and discuss possible causes. (c) Fit both a linear and an exponential trend. (c) Interpret each fitted trend equation, explaining the implications. (d) Make a projection for 2005. Do you believe it? (e) To whom is this issue relevant? *Note:* Time increments are 5 years, so use $t = 10$ for the 2005 forecast. Petroleum

U.S. Petroleum Imports, Exports, and Consumption (million barrels per day)

Year	Imports	Exports	Net Imports	Total	% of World
1960	1.810	0.200	1.610	9.800	45.9
1965	2.470	0.190	2.280	11.510	37.0
1970	3.420	0.260	3.160	14.700	31.4
1975	6.056	0.209	5.846	16.322	29.0
1980	6.909	0.544	6.365	17.056	27.0
1985	5.067	0.781	4.286	15.726	26.2
1990	8.018	0.857	7.161	16.988	25.7
1995	8.840	0.949	7.886	17.720	25.3
2000	11.460	1.040	10.419	19.701	25.6

Source: U.S. Dept. of Transportation, www.bts.gov.

14.28 (a) Choose *one* prison time series and plot it on a graph. (b) Describe the trend (if any) and discuss possible causes. (c) Fit both a linear and an exponential trend. (d) Interpret each fitted trend equation, explaining its implications. (e) Using both models, make a projection for 2010. Do you believe it? Explain. **Prisoners**

U.S. Adults on Probation, in Jail or Prison, or on Parole: 1986–2002 (thousands)

Year	Total	% of Adult Pop	Probation	Jail	Prison	Parole
1986	3,239	1.8	2,115	273	526	326
1987	3,460	1.9	2,247	294	563	356
1988	3,714	2.0	2,356	342	608	408
1989	4,056	2.2	2,522	393	683	457
1990	4,348	2.3	2,670	403	743	531
1991	4,536	2.4	2,728	424	793	590
1992	4,763	2.5	2,812	442	851	659
1993	4,944	2.6	2,903	456	909	676
1994	5,141	2.7	2,981	480	990	690
1995	5,335	2.8	3,078	499	1,079	679
1996	5,483	2.8	3,165	510	1,128	680
1997	5,726	2.9	3,297	558	1,177	695
1998	6,126	3.1	3,670	584	1,224	696
1999	6,331	3.1	3,780	596	1,287	714
2000	6,437	3.1	3,826	614	1,316	724
2001	6,574	3.1	3,932	624	1,330	732
2002	6,684	3.1	3,955	658	1,368	753

Source: *Statistical Abstract of the United States, 2004.*

14.29 (a) Choose *two* time series on SAT scores that you would like to compare. Plot both series on the same graph. (b) Fit a linear trend to each series. (c) Interpret each fitted trend equation. (d) What are the implications (if any) of your analysis, and for whom? **SAT**

SAT Averages for College-Bound H.S. Seniors, 1990–2001

	Verbal Score			Mathematical Score		
Year	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*
1990–91	499	503	495	500	520	482
1991–92	500	504	496	501	521	484
1992–93	500	504	497	503	524	484
1993–94	499	501	497	504	523	487
1994–95	504	505	502	506	525	490
1995–96	505	507	503	508	527	492
1996–97	505	507	503	511	530	494
1997–98	505	509	502	512	531	496
1998–99	505	509	502	511	531	495
1999–00	505	507	504	514	533	498
2000–01	506	509	502	514	533	498

Source: College Entrance Examination Board, National Report on College-Bound Seniors, various years. www.collegeboard.com.

14.30 (a) Use Excel, MegaStat, or MINITAB to fit an m-period moving average to the exchange rate data shown below with $m = 2, 3, 4,$ and 5 periods. Make a line chart. (b) Which value of m do you prefer? Why? (c) Is a moving average appropriate for this kind of data? Include a chart for each value of m. **Sterling**

Daily Spot Exchange Rate, U.S. Dollars per Pound Sterling

Date	*Rate*	*Date*	*Rate*	*Date*	*Rate*	*Date*	*Rate*
1-Apr-04	1.8564	16-Apr-04	1.8004	3-May-04	1.7720	18-May-04	1.7695
2-Apr-04	1.8293	19-Apr-04	1.8055	4-May-04	1.7907	19-May-04	1.7827
5-Apr-04	1.8140	20-Apr-04	1.7914	5-May-04	1.7932	20-May-04	1.7710
6-Apr-04	1.8374	21-Apr-04	1.7720	6-May-04	1.7941	21-May-04	1.7880
7-Apr-04	1.8410	22-Apr-04	1.7684	7-May-04	1.7842	24-May-04	1.7908
8-Apr-04	1.8325	23-Apr-04	1.7674	10-May-04	1.7723	25-May-04	1.8135
9-Apr-04	1.8322	26-Apr-04	1.7857	11-May-04	1.7544	26-May-04	1.8142
12-Apr-04	1.8358	27-Apr-04	1.7925	12-May-04	1.7743	27-May-04	1.8369
13-Apr-04	1.8160	28-Apr-04	1.7720	13-May-04	1.7584	28-May-04	1.8330
14-Apr-04	1.7902	29-Apr-04	1.7751	14-May-04	1.7572		
15-Apr-04	1.7785	30-Apr-04	1.7744	17-May-04	1.7695		

Source: Federal Reserve Board of Governors.

14.31 Refer to exercise 14.30. (a) Plot the dollar/pound exchange rate data. Make the graph nice, then copy and paste it so you have four copies. (b) Use MegaStat or MINITAB to perform a simple exponential smoothing using $\alpha = .05, .10, .20,$ and $.50$, using a different line chart for each. (c) Which value of α do you prefer? Why? (d) Is an exponential smoothing process appropriate for this kind of data? **Sterling**

14.32 (a) Plot the data on gas bills. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. (d) Which months are the most expensive? The least expensive? Can you explain this pattern? (e) Is there a trend in the deseasonalized data? *(f) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. **GasBills**

Natural Gas Bills for a Residence, 2000–2003

Month	*2000*	*2001*	*2002*	*2003*
Jan	78.98	118.86	101.44	155.37
Feb	84.44	111.31	122.20	148.77
Mar	65.54	75.62	99.49	115.12
Apr	62.60	77.47	55.85	85.89
May	29.24	29.23	44.94	46.84
Jun	18.10	17.10	19.57	24.93
Jul	91.57	16.59	15.98	20.84
Aug	6.48	27.64	14.97	26.94
Sep	19.35	28.86	18.03	34.17
Oct	29.02	48.21	56.98	88.58
Nov	94.09	67.15	115.27	100.63
Dec	101.65	125.18	130.95	174.63

Source: Homeowner's records.

14.33 (a) Plot the data on building permits. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. (d) Which months have the most permits? The fewest? Is this logical? (e) Is there a trend in the deseasonalized data? **Permits**

Single Family Home Building Permits in Southeastern Michigan

Month	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*
Jan	763	981	986	999	830	1,155
Feb	877	1,058	1,146	1,129	1,029	1,138
Mar	1,330	1,448	1,384	1,705	1,716	1,779
Apr	1,530	2,080	2,100	1,817	1,845	1,670
May	1,719	2,036	1,699	1,762	1,909	1,692
Jun	1,787	1,723	1,643	1,955	2,037	1,634
Jul	1,440	1,869	1,605	1,746	1,841	1,414
Aug	1,790	1,737	1,635	1,476	1,885	1,614
Sep	1,529	1,502	1,593	1,625	1,584	1,418
Oct	1,536	1,767	1,672	1,720	1,643	1,618
Nov	1,346	1,217	1,059	1,530	1,296	1,173
Dec	938	1,050	1,111	1,367	1,158	693

Source: www.semcog.org.

14.34 (a) Plot the data on airplane shipments. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. Is there a trend in the deseasonalized data? **AirplanesQtr**

U.S. Manufactured General Aviation Shipments, 1986–2003

Year	Qtr 1	Qtr 2	Qtr 3	Qtr 4	Total
1986	285	364	393	453	1,495
1987	227	330	239	289	1,085
1988	260	291	252	340	1,143
1989	304	361	425	445	1,535
1990	269	294	274	297	1,144
1991	250	262	237	272	1,021
1992	193	200	238	225	941
1993	170	194	246	260	964
1994	181	225	209	266	928
1995	208	248	257	315	1,077
1996	229	284	230	310	1,115
1997	253	337	367	525	1,549
1998	481	486	546	602	2,200
1999	502	611	606	702	2,504
2000	613	704	685	712	2,816
2001	568	711	586	673	2,632
2002	442	576	510	641	2,207
2003	393	526	492	679	2,137

Note: Quarterly shipments may not add to annual total because some manufacturers report only annual totals.

Source: U.S. Manufactured General Aviation Shipments, *Statistical Databook 2003*, General Aviation Manufacturers Association, used with permission.

14.35 (a) Plot the data on revolving credit (credit cards and home equity lines of credit are the two major types of revolving credit). (b) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. Is there a trend in the deseasonalized data? (c) Which months have the most borrowing? The least? Is this logical? **Revolving**

U.S. Consumers Revolving Credit (billions)

Month	2001	2002	2003	2004
Jan	223.2	232.5	240.6	276.7
Feb	221.5	229.7	239.7	272.8
Mar	220.1	230.2	234.0	268.3
Apr	227.7	235.6	235.4	270.6
May	229.1	233.1	240.4	278.0
Jun	225.7	231.0	240.7	275.6
Jul	222.1	229.9	238.6	278.7
Aug	219.6	241.1	240.7	286.4
Sep	216.3	243.1	239.9	286.7
Oct	223.3	242.4	235.8	286.1
Nov	233.2	244.2	269.5	285.8
Dec	238.3	250.2	284.7	315.8

Source: www.federalreserve.gov.

14.36 (a) Plot the data on jewelry sales. (b) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. Is there a trend in the deseasonalized data? (c) Which months have the most sales? The least? Is this logical? *(d) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. **Jewelry**

U.S. Jewelry Sales, 1990–1995 ($ millions)

Month	1990	1991	1992	1993	1994	1995
Jan	846	821	813	801	897	921
Feb	1,025	998	1,042	1,001	1,181	1,230
Mar	984	967	930	901	1,048	1,145
Apr	1,004	1,012	985	1,005	1,159	1,213
May	1,263	1,313	1,190	1,244	1,354	1,616
Jun	1,134	1,099	1,111	1,268	1,244	1,402
Jul	1,075	1,021	1,051	1,277	1,213	1,272
Aug	1,132	1,058	1,103	1,268	1,308	1,408
Sep	996	963	1,046	1,188	1,234	1,340
Oct	1,084	1,080	1,135	1,210	1,313	1,387
Nov	1,400	1,329	1,378	1,557	1,717	1,891
Dec	3,238	3,071	3,475	3,822	4,171	4,526

Source: www.census.gov, project by statistics student Devora Davis.

14.37 (a) Plot the data on M1 money stock. (b) Use MegaStat or MINITAB to calculate estimated seasonal indexes and trend. Is there a trend in the deseasonalized data? (c) Make monthly forecasts for 2002. *Note:* M1 includes currency, travelers checks, demand deposits, and other checkable deposits. **MoneyStock**

U.S. Money Stock M1 Component, 1995–2001 ($ billions)

Month	1995	1996	1997	1998	1999	2000	2001
Jan	1,159.0	1,129.4	1,086.3	1,079.4	1,103.0	1,126.4	1,099.6
Feb	1,134.9	1,105.1	1,065.2	1,065.5	1,084.3	1,096.8	1,087.5
Mar	1,138.9	1,117.3	1,067.5	1,075.2	1,096.6	1,108.1	1,107.4
Apr	1,159.9	1,131.1	1,073.1	1,086.9	1,112.6	1,124.9	1,122.7
May	1,133.5	1,105.1	1,053.6	1,070.0	1,095.4	1,100.4	1,111.0
Jun	1,140.4	1,114.2	1,064.1	1,074.3	1,097.2	1,102.6	1,122.6
Jul	1,145.2	1,110.0	1,065.3	1,073.5	1,096.7	1,104.0	1,135.9
Aug	1,138.9	1,097.0	1,068.7	1,068.6	1,092.7	1,095.9	1,141.3
Sep	1,138.0	1,091.2	1,059.1	1,070.0	1,086.3	1,090.5	1,194.3
Oct	1,132.4	1,077.6	1,057.1	1,076.8	1,095.3	1,093.6	1,155.5
Nov	1,138.0	1,086.4	1,073.6	1,097.3	1,113.3	1,093.3	1,164.8
Dec	1,152.1	1,104.7	1,096.9	1,120.4	1,148.3	1,112.3	1,202.2

Source: Federal Reserve Board of Governors, www.federalreserve.gov. Data not seasonally adjusted.

14.38 (a) Use MegaStat or MINITAB to deseasonalize the quarterly data on Coca-Cola's revenues and fit a trend. Interpret the results. (b) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. (c) Use the regression equation to make a prediction for each quarter in 2005. *(d) If you have access to *Standard & Poor's Stock Reports, 2006,* check your forecasts. How accurate were they? **CocaCola**

Coca-Cola Revenues ($ millions), 1999–2004

Quarter	1999	2000	2001	2002	2003	2004
1	4,428	4,391	4,479	4,079	4,502	5,078
2	5,379	5,621	5,293	5,368	5,695	5,965
3	5,195	5,543	5,397	5,322	5,671	5,662
4	4,931	4,903	4,923	4,795	5,176	5,257
Year	19,933	20,458	20,092	19,564	21,044	21,962

Source: *Standard & Poor's Stock Reports,* March 2005.

14.39 (a) Use MegaStat or MINITAB to perform a regression using seasonal binaries. Interpret the results. (b) Make monthly forecasts for 1997. If you can find data on the Web, check your forecasts.
StudentPilots

Student Pilot Certificates Issued By Month, 1991–1996

Month	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*
Jan	5,704	6,104	4,802	4,394	4,505	4,067
Feb	5,541	5,773	5,144	4,562	4,307	4,057
Mar	5,950	6,773	5,835	5,696	5,189	4,301
Apr	6,513	6,703	5,507	5,308	4,744	4,758
May	6,622	6,299	5,597	5,788	5,396	5,065
Jun	7,932	7,819	6,683	6,837	5,878	5,031
Jul	8,442	8,074	6,758	6,011	5,708	5,807
Aug	8,580	7,210	7,191	7,054	6,590	5,564
Sep	7,630	7,251	6,343	6,274	6,001	5,192
Oct	7,956	6,760	5,797	5,790	4,000	5,310
Nov	7,661	5,240	5,117	4,785	4,179	4,240
Dec	3,674	4,371	4,404	4,002	4,000	3,261

Source: Federal Aviation Administration, http://api.hq.faa.gov/handbook/1996.

*14.40 Translate each of the following fitted exponential trend models into a compound interest model of the form $y_t = y_0(1+r)^t$. *Hint:* See *LearningStats* Unit 14 or footnote on p. 615.

a. $y_t = 456e^{.123t}$ b. $y_t = 228e^{.075t}$ c. $y_t = 456e^{-.038t}$

*14.41 Translate each of the following fitted compound interest trend models into an exponential model of the form $y_t = ae^{bt}$. *Hint:* See *LearningStats* Unit 14 or footnote on p. 615.

a. $y_t = 123(1.089)^t$ b. $y_t = 654(1.217)^t$ c. $y_t = 308(.942)^t$

Related Reading

Brocklebank, John C.; and David A. Dickey. *SAS for Forecasting Time-Series*. Wiley, 2003.

DeLurgio, Stephen A. *Forecasting Principles and Applications*. Irwin/McGraw-Hill, 1998.

Diebold, Francis X. *Elements of Forecasting*. 3rd ed. South-Western, 2004.

Gaynor, Patricia E.; and Rickey C. Kirkpatrick. *Introduction to Time-Series Modeling and Forecasting in Business and Economics*. McGraw-Hill, 1994.

Granger, C. W. J. *Forecasting in Business and Economics*. 2nd ed. Academic Press, 1989.

Hanke, John E.; and Dean W. Wichern. *Business Forecasting*. 8th ed. Prentice-Hall, 2005.

Pindyck, Robert S.; and Daniel L. Rubinfeld. *Econometric Models and Economic Forecasts*. 4th ed. Irwin/McGraw-Hill, 1998.

Wilson, J. Holton; and Barry Keating. *Business Forecasting*. 4th ed. Irwin, 2002.

Makridakis, Spyros; Steven C. Wheelwright; and Rob J. Hyndman. *Forecasting: Methods and Applications*. 4th ed. Wiley, 2006.

LearningStats Unit 14 Time-Series Analysis

LS

LearningStats Unit 14 contains simulations to generate quarterly and monthly data, case studies of various kinds of fitted trends, illustrations of the components of a time series, and examples of student projects using fitted trends. Modules are designed for self-study, so you can concentrate on material that is new, and pass quickly over things that you already know or do not find interesting. Your instructor may assign specific modules, or you may decide to check them out because the topic sounds interesting. In addition to helping you learn about statistics, they may be useful as references later on.

Topic	*LearningStats Modules*
Trends and forecasting	Time-Series Components Trend Fitting Trend Forecasting
Using Excel	Excel Trends—1 Excel Trends—2 Exponential Trend Dissimilar Magnitudes Exponential Smoothing
Simulations	Time-Series Components Trend Simulator Seasonal Time-Series Generator
Fit and seasonality	Trend Fit Measures Seasonal Factors
Case studies	Health Trends Olympic Times Federal Budget Male/Female Income Brad's Bowling Scores
Student projects	Gas Prices Investing Olympic Times

Key: = PowerPoint = Word = Excel

Visual Statistics

VS

Visual Statistics Modules on Time-Series

Module	*Module Name*
20	Visualizing Time-Series Data

Visual Statistics Module 20 is designed to help you

- Understand the importance of the data collection period.
- Recognize the difficulty in separating trend, seasonality.
- See why sample size is important.
- Understand the difference between additive and multiplicative seasonality.
- Understand common statistics of fit (*MAPE*, R^2, standard error).

The worktext (included on the CD in .PDF format) contains lists of concepts covered, objectives of the modules, overviews of concepts, illustrations of concepts, orientations to module features, learning exercises (basic, intermediate, advanced), learning projects (individual, team), self-evaluation quizzes, glossaries of terms, and solutions to self-evaluation quizzes.

CHAPTER

14 Time-Series Analysis

CHAPTER CONTENTS

CHAPTER LEARNING OBJECTIVES

When you finish this chapter, you should be able to

LO 14-1 Define time-series data and its components.

LO 14-2 Interpret a linear, exponential, or quadratic trend model.

LO 14-3 Fit any common trend model and use it to make forecasts.

LO 14-4 Know the definitions of common fit measures.

LO 14-5 Interpret a moving average and use Excel to create it.

LO 14-6 Use exponential smoothing to forecast trendless data.

LO 14-7 Interpret seasonal factors and use them to make forecasts.

LO 14-8 Use regression with seasonal binaries to make forecasts.

LO 14-9 Interpret index numbers.

JUNG YEON-JE/AFP/Getty Images

14.1 TIME-SERIES COMPONENTS

Time-Series Data

Businesses must track their performance. By looking at their sales, costs, or profits over time, businesses can tell where they've been, whether they are performing poorly or satisfactorily, and how much improvement is needed, in both the short term and the long term. A **time-series variable** (denoted *Y*) consists of data observed over *n* periods of time. Consider a clothing retailer that specializes in blue jeans. Examples of time-series data this company might be interested in tracking would be the number of jeans sold and the company's market share. Or, from the manufacturing perspective, the company might track cost of raw materials over time.

LO 14.1

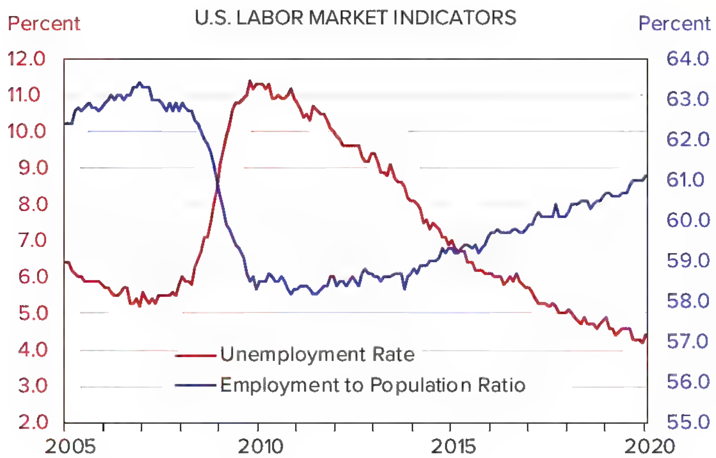
Define time-series data and its components.

Businesses also use time-series data to monitor whether a particular process is stable or unstable. And they use time-series data to help anticipate the future, a process we call *forecasting*. In addition to business time-series data, we see economic time-series data in *The Wall Street Journal* or *Bloomberg Businessweek* and also in *USA Today* or *Time* or even when we browse the web. Although business and economic time-series data are most common, we can see time-series data for population, health, crime, sports, and social problems. Usually, time-series data are presented in a graph, like Figures 14.1 and 14.2.

It is customary to plot time-series data either as a line graph or as a bar graph, with time on the horizontal *X*-axis and the variable of interest on the vertical *Y*-axis to reveal how the variable changes over time. In a line graph, the *X*-*Y* data points are connected with line segments to make it easier to see fluctuations. While anyone can understand time-series graphs in a general way, this chapter explains how to interpret time-series data *statistically* and to make

Figure 14.1

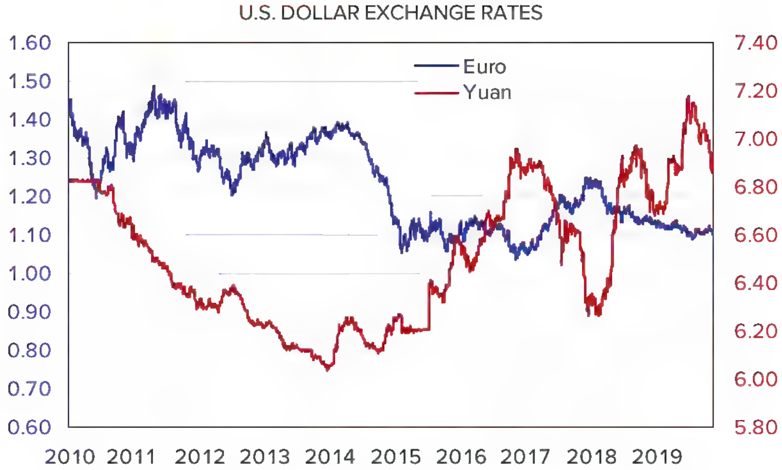
U.S. Employment (monthly, not seasonally adjusted) **Labor**



Source: www.bls.gov. Latest data shown are for January 1, 2020.

Figure 14.2

Exchange Rates (daily) Exchange



Source: www.federalreserve.gov. Latest data shown are for January 24, 2020.

defensible forecasts. Our analysis begins with sample observations $y_1, y_2, \ldots, y_n$ covering n time periods. The following notation is used:

- y_t is the value of the time series in period t.
- t is an index denoting the time period ($t = 1, 2, \ldots, n$).
- n is the number of time periods.
- $y_1, y_2, \ldots, y_n$ is the data set for analysis.

To distinguish time-series data from cross-sectional data, we use y_t for an individual observation, with a subscript t instead of i.

Time-series data may be measured *at a point in time* or *over an interval of time.* For example, in accounting, balance sheet data are measured at the end of the fiscal year, while income statement data are measured over an entire fiscal year. The gross domestic product (GDP) is a flow of goods and services measured *over an interval of time,* while the prime rate of interest is measured *at a point in time.* Your GPA is measured *at a point in time,* while your weekly pay is measured *over an interval of time.* The distinction is sometimes vague in reported data, but a little thought will usually clarify matters. For example, Canada's 2018 unemployment rate (5.6 percent) would be measured at a point in time (e.g., at year's end), while Canada's 2018 hydroelectric production (382 terawatt-hours) would be measured over the entire year (see www.statcan.gc.ca).

Periodicity

The **periodicity** is the time interval over which data are collected (decade, year, quarter, month, week, day, hour). For example, the U.S. population is measured each *decade,* your personal income tax is calculated *annually,* GDP is reported *quarterly,* the unemployment rate is estimated *monthly,* and *The Wall Street Journal* reports the closing price of Apple stock *daily* (although stock prices are also monitored continuously on the web). Firms typically report profits by quarter but pension liabilities only at year's end. Any periodicity is possible, but the principles of time-series modeling can be understood with three common data types:

- Annual data (1 observation per year)
- Quarterly data (4 observations per year)
- Monthly data (12 observations per year)

Time-Series Components

Time-series *decomposition* seeks to separate a time-series *Y* into four components: trend (*T*), cycle (*C*), seasonal (*S*), and irregular (*I*). Figure 14.3 illustrates these four components in a hypothetical monthly time series. The four components may be thought of as layering atop one another to produce the actual time series. In this example, the irregular component (*I*) is large enough to obscure the cycle (*C*) and seasonal (*S*) components but not the trend (*T*). However, we can usually extract the original components from the time series by using statistical methods. These components are assumed to follow either an **additive model** or a **multiplicative model**, as shown in Table 14.1.

Figure 14.3

Four Components of a Time Series

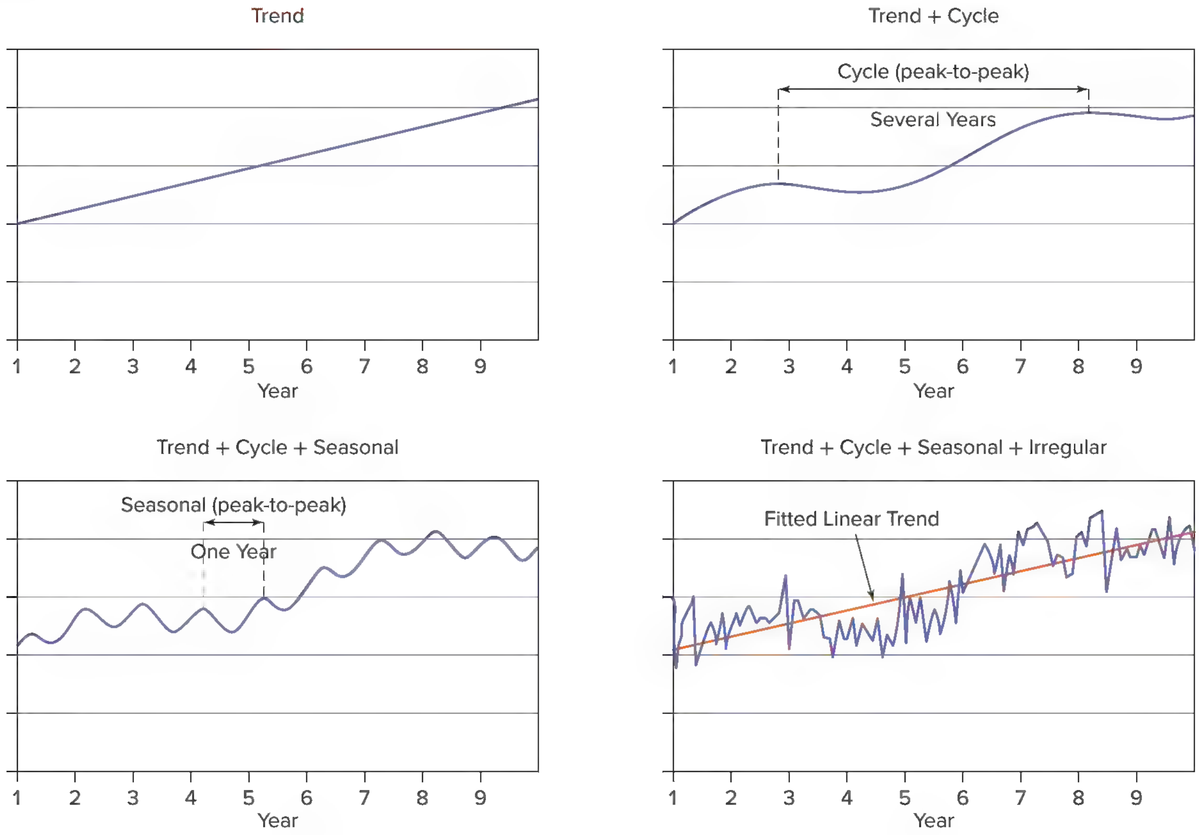


Table 14.1

Components of a Time Series

Model	Components	Used for
Additive	$Y = T + C + S + I$	Data of similar magnitude (short-run or trend-free data) with constant *absolute* growth or decline.
Multiplicative	$Y = T \times C \times S \times I$	Data of increasing or decreasing magnitude (long-run or trended data) with constant *percent* growth or decline.

The additive form is attractive for its simplicity, but the multiplicative model is often more useful for forecasting financial data, particularly when the data vary over a range of magnitudes. Especially in the short run, it may not matter greatly which form is assumed. In fact, the model forms are fundamentally equivalent because the multiplicative model becomes additive if logarithms are taken (as long as the data are nonnegative):

$$\log(Y) = \log(T \times C \times S \times I) = \log(T) + \log(C) + \log(S) + \log(I)$$

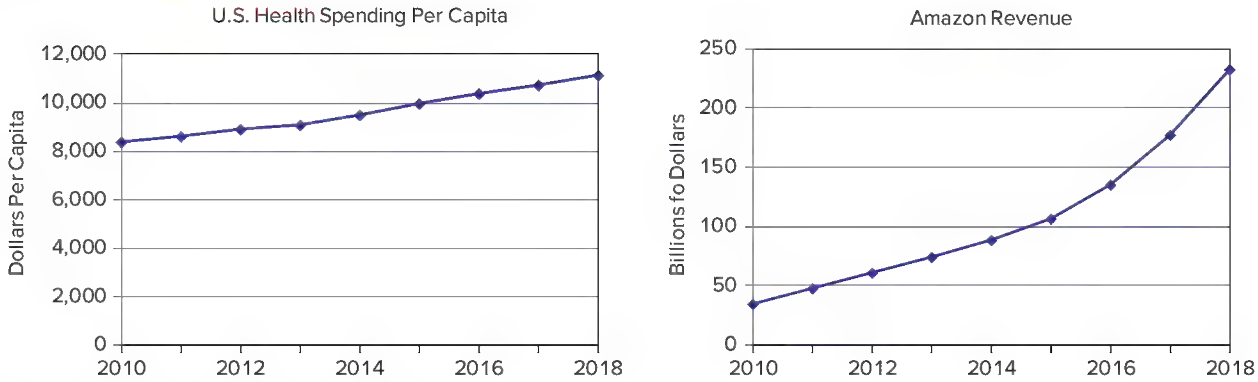
Trend

Trend (T) is a general movement over all years ($t = 1, 2, \ldots, n$). Change over a few years is not a trend. Some trends are steady and predictable. For example, the data may be steadily growing (e.g., total U.S. population), neither growing nor declining (e.g., your current car's mpg), or steadily declining (infant mortality rates in a developing nation). A mathematical trend can be fitted to any data, but its predictive value depends on the situation. For example, to predict health expenditures or Amazon's net sales (Figure 14.4), a mathematical trend might be useful, but a mathematical model might not be very helpful for predicting frequency of hurricanes or Fargo, ND, snowfall (Figure 14.5).

Most of us think of three general patterns: growth, stability, or decline. But there are subtler trends within each category. A time series can increase at a steady *linear* rate (e.g., the

Figure 14.4

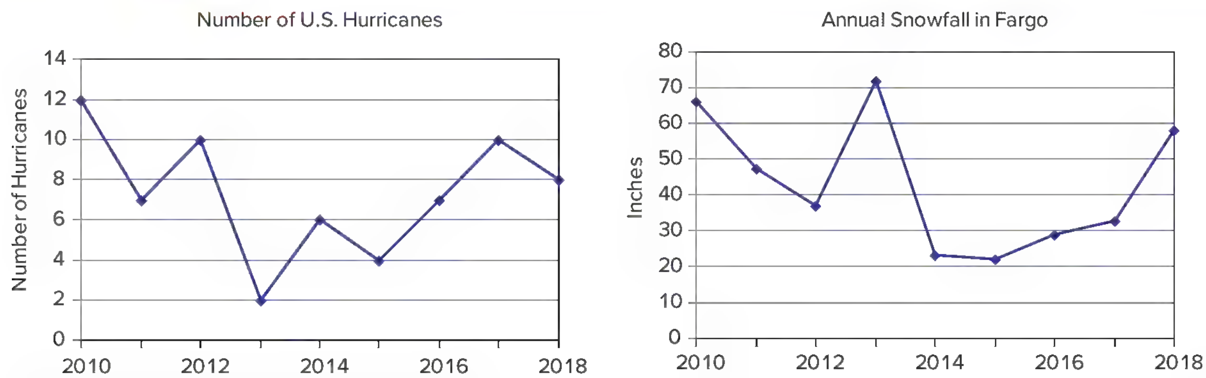
Steady Trend Steady



Sources: www.healthsystemtracker.org and Mergent Online.

Figure 14.5

Erratic Pattern Erratic



Sources: www.nhc.noaa.gov/ and https://www.ncdc.noaa.gov/.

number of books you have read in your lifetime), at an *increasing* rate (e.g., Medicare costs for an aging population), or at a *decreasing* rate (e.g., live attendance at NFL football games). It can grow for a while and then level off (e.g., sales of 86-inch TVs) or grow toward an asymptote (e.g., percent of adults owning an iPhone).

Cycle

Cycle (C) is a repetitive up-and-down movement around the trend that covers *several years.* For example, industry analysts have studied cycles for sales of new automobiles, new home construction, inventories, and business investment. These cycles are based primarily on product life and replacement cycles. In any market economy, there are broad business cycles that affect employment and production. After we have extracted the trend and seasonal components of a time series, a cycle may be detected as autocorrelation in the residuals (see Chapter 12, Section 12.8). Although cycles are conceptually important, there is no general theory of cycles, and even those cycles that have been identified in specific industries have erratic timing and complex causes that defy generalization. Over a small number of time periods (a typical forecasting situation), cycles are undetectable or may resemble a trend. For this reason cycles are not discussed further in this chapter.

Seasonal

Seasonal (S) is a repetitive cyclical pattern *within a year.** For example, many retail businesses experience strong sales during the fourth quarter because of Christmas. Automobile sales rise when new models are released. Peak demand for airline flights to Europe occurs during summer vacation travel. Although often imagined as sine waves, seasonal patterns may not be smooth. Peaks and valleys can occur in any month or quarter, and each industry may face its own unique seasonal pattern. For example, June weddings tend to create a "spike" in bridal sales, but there is no "sine wave" pattern in bridal sales. By definition, annual data have no seasonality.

Irregular

Irregular (I) is a random disturbance that follows no apparent pattern. It also is called the *error* component or *random noise* reflecting all factors other than trend, cycle, and seasonality. For example, daily prices of many common stocks fluctuate greatly. When the irregular component is large, it may be difficult to isolate other individual model components. In such cases, we use special techniques (e.g., **moving average** or **exponential smoothing**) to make short-run forecasts. Faced with erratic data, experts may use their own knowledge to make *judgment forecasts.* For example, vehicle sales forecasts may combine judgment forecasts from dealers, financial staff, and economists. However, a major systemic shock such as the COVID-19 crisis may dominate other time series components and render forecasting efforts moot.

14.2 TREND FORECASTING

LO 14-2

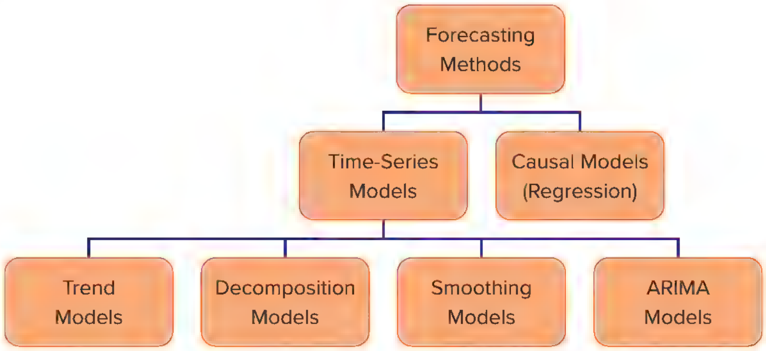
Interpret a linear, exponential, or quadratic trend model.

There are many forecasting methods designed for specific situations. Much of this chapter deals with *trend models* because they are so common in business. You also will learn to use *decomposition* to make adjustments for *seasonality* and how to use *smoothing models.* The important topics of *ARIMA models* and *causal models* are reserved for a more specialized class in forecasting. Figure 14.6 summarizes the main categories of forecasting models.

*Repetitive patterns within a week, day, or other time period also may be considered seasonal. For example, mail volume in the U.S. Postal Service is higher on Monday. Emergency arrivals at hospitals are lower during the first shift (between midnight and 6:00 a.m.). In this chapter, we will discuss only *monthly* and *quarterly* seasonal patterns because these are most typical of business data.

Figure 14.6

Overview of Forecasting



Three Trend Models

There are many possible trend models, but three of them are especially useful in business:

(14.1) $y_t = a + bt$ for $t = 1, 2, \dots, n$ (linear trend)

(14.2) $y_t = ae^{bt}$ for $t = 1, 2, \dots, n$ (exponential trend)

(14.3) $y_t = a + bt + ct^2$ for $t = 1, 2, \dots, n$ (quadratic trend)

The linear and exponential models are widely used because they have only two parameters and are familiar to most business audiences. The quadratic model may be useful when the data have a turning point. All three can be fitted by Excel.

Linear Trend Model

The **linear trend** model has the form $y_t = a + bt$. It is useful for a time series that grows or declines by the same amount (b) in each period, as shown in Figure 14.7. It is the simplest model and may suffice for short-run forecasting. It is generally preferred in business as a baseline forecasting model unless there are compelling reasons to consider a more complex model.

Figure 14.7

Linear Trend Models

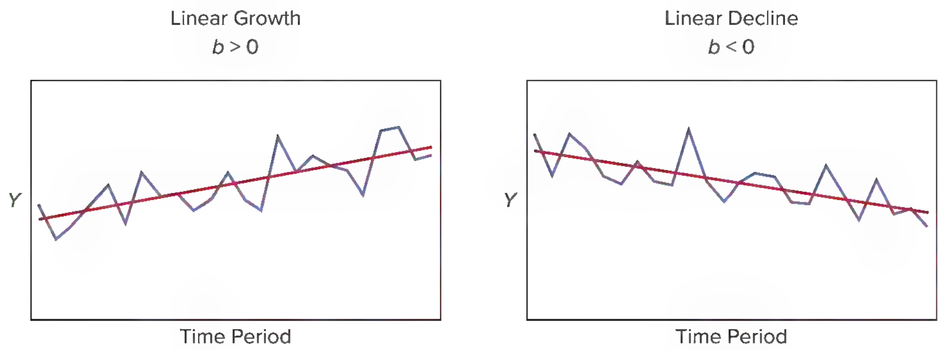


Illustration: Linear Trend

Following the population "echo boom," some U.S. states (especially in the Northeast and Midwest) began to see declines in the number of high school graduates. Concerned about its potential loss of traditional student populations aged 18–25, a Midwestern university wanted to extrapolate the recent trend in fall semester enrollments. The slope of Excel's fitted trend, shown in Figure 14.8, indicates that, on average, the university has lost 235 students per year.

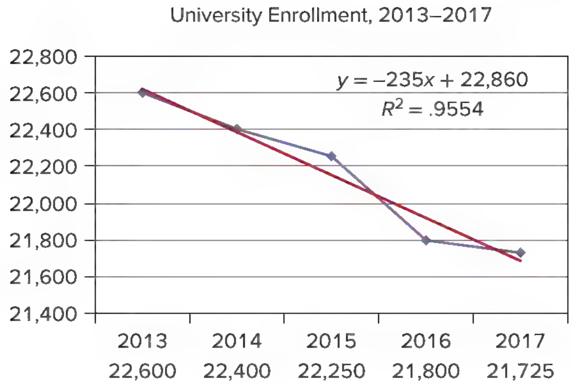


Figure 14.8

Excel's Linear Trend
Enrollment

Linear Trend Calculations

LO 14-3

Fit any common trend model and use it to make forecasts.

The linear trend is fitted in the usual way by using the ordinary least squares formulas, as illustrated in Table 14.2. Because you are already familiar with regression, we will only point out the use of the index $t = 1, 2, 3, 4, 5$ as the independent variable (instead of using the years 2013, 2014, 2015, 2016, 2017). We use this time index to simplify the calculations and keep the data magnitudes under control (Excel uses this method too).

$$\textit{Slope:} \qquad b = \frac{\sum_{t=1}^{n}(t-\bar{t})(y_t-\bar{y})}{\sum_{t=1}^{n}(t-\bar{t})^2} = \frac{-2{,}350}{10} = -235$$

$$\textit{Intercept:} \qquad a = \bar{y} - b\bar{t} = 22{,}155 - (-235)(3) = 22{,}860$$

The *slope* of the fitted trend $y_t = 22{,}860 - 235t$ says that, unless the university takes steps to recruit new, nontraditional student populations, it can expect to lose 235 students each year ($dy_t/dt = -235$). The *intercept* is the "starting point" for the time series in period $t = 0$; that is, $y_0 = 22{,}860 - 235(0) = 22{,}860$.

Table 14.2

Sums for Least Squares Calculations

Year	t	y_t	$t-\bar{t}$	$y_t-\bar{y}$	$(t-\bar{t})^2$	$(t-\bar{t})(y_t-\bar{y})$
2013	1	22,600	−2	445	4	−890
2014	2	22,400	−1	245	1	−245
2015	3	22,250	0	95	0	0
2016	4	21,800	1	−355	1	−355
2017	5	21,725	2	−430	4	−860
Sum	15	110,775	0	0	10	−2,350
Mean	3	22,155	0	0	2	−470

Fitting and Interpreting an Annual Trend

In fitting a trend to annual data, the years (2013, 2014, 2015, 2016, 2017) are merely used as labels for the *X*-axis. The yearly labels should *not* be used in fitting the trend or calculating the forecast. To fit a trend to annual data, convert the labels to a time index ($t = 1, 2, \ldots,$ etc.). To make a forecast, insert a value for the time index ($t = 1, 2, \ldots,$ etc.) into Excel's fitted trend.

Forecasting a Linear Trend

We can make a forecast for any future year by using the fitted model $y_t = 22{,}860 - 235t$. In the enrollment example, the fitted trend equation is based on only 5 years' data, so we should be wary of extrapolating very far ahead:

$$\text{For } 2018\,(t = 6)\text{: } y_6 = 22{,}860 - 235(6) = 21{,}450$$
$$\text{For } 2019\,(t = 7)\text{: } y_7 = 22{,}860 - 235(7) = 21{,}215$$
$$\text{For } 2020\,(t = 8)\text{: } y_8 = 22{,}860 - 235(8) = 20{,}980$$

Linear Trend: Calculating R^2

The worksheet shown in Table 14.3 shows the calculation of the coefficient of determination. In this illustration, the linear model gives a good fit R^2 = .9554) to the *past* data. However, a good fit to the past data does not guarantee good *future* forecasts. A deeper analysis of underlying causes of enrollment declines is needed. Are the causal forces likely to remain the same in subsequent years? Could the current demographic decline continue indefinitely, or will enrollments approach an asymptote or even start to grow again? These are questions that forecasters must ask. The forecast is simply a projection of current trend assuming that nothing changes.

$$\textit{Coefficient of determination: } R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = 1 - \frac{25{,}750}{578{,}000} = .9554$$

Table 14.3

Sums for R^2 Calculations

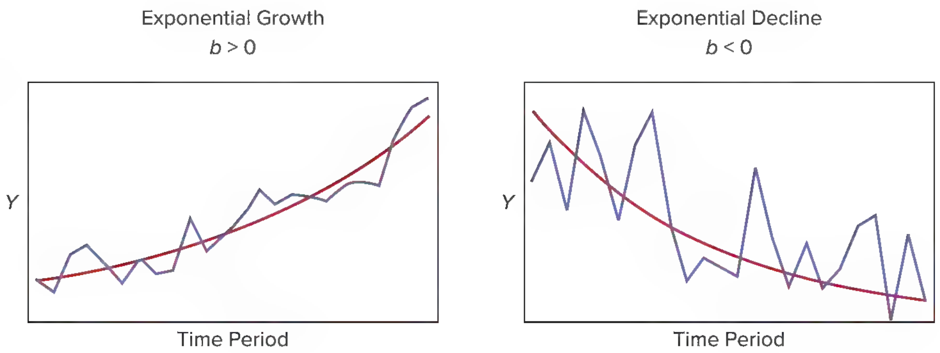
Year	t	y_t	$\hat{y}_t = 22{,}860 - 235t$	$y_t - \hat{y}_t$	$(y_t - \hat{y}_t)^2$	$(y_t - \bar{y})^2$
2013	1	22,600	22,625	−25	625	198,025
2014	2	22,400	22,390	10	100	60,025
2015	3	22,250	22,155	95	9,025	9,025
2016	4	21,800	21,920	−120	14,400	126,025
2017	5	21,725	21,685	40	1,600	184,900
Sum	15		110,775	0	25,750	578,000

Exponential Trend Model

The **exponential trend** model has the form $y_t = ae^{bt}$. It is useful for a time series that grows or declines at the same *rate* (b) in each period, as shown in Figure 14.9. When the growth rate is positive ($b > 0$), then Y grows by an *increasing* amount each period (unlike the linear model, which assumes a *constant* increment each period). If the growth rate is negative ($b < 0$), then Y declines by a *decreasing* amount each period (unlike the linear model, which assumes a *constant* decrement each period).

Figure 14.9

Exponential Trend Models



When to Use the Exponential Model

The exponential model is often preferred for financial data or data that cover a longer period of time. When you invest money in a commercial bank savings account, interest accrues at a given percent. Your savings grow faster than a linear rate because you earn interest on the accumulated interest. Banks use the exponential formula to calculate interest on CDs. Financial analysts often find the exponential model attractive because costs, revenue, and salaries are best projected under assumed *percent* growth rates.

Another nice feature of the exponential model is that you can compare two growth rates in two time-series variables with dissimilar data units (i.e., a percent growth rate is *unit-free*). For example, between 2000 and 2018 the number of Medicare enrollees grew from 40.0 million persons to 59.9 million persons (2.3 percent annual growth rate), while Medicare payments grew from \$217 billion to \$731 billion (7.0 percent annual growth rate). Comparing these percents, we see that Medicare insurance payments have been growing three times as fast as the Medicare head count (see www.cms.gov). These facts underlie the ongoing debate about Medicare spending in the United States.

There may not be much difference between a linear and exponential model when the growth rate is small and the data set covers only a few time periods. For example, suppose your starting salary is \$50,000. Table 14.4 compares salary increases of \$2,500 each year ($y_t = 50{,}000 + 2{,}500t$) with a continuously compounded 4.879 percent salary growth ($y_t = 50{,}000e^{.04879t}$). Over the first few years, there is little difference. But after 20 years, the difference is obvious, as shown in Figure 14.10. Despite its attractive simplicity,* the linear model's assumptions may be inappropriate for some financial variables.

Table 14.4

Two Models of Salary Growth

t	$y_t = 50{,}000 + 2{,}500t$ Linear	$y_t = 50{,}000e^{.04879t}$ Exponential
0	50,000	50,000
5	62,500	63,814
10	75,000	81,445
15	87,500	103,946
20	100,000	132,665

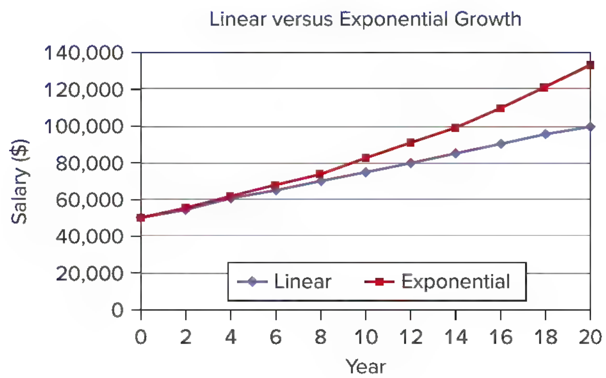


Figure 14.10

Linear and Exponential Growth Compared

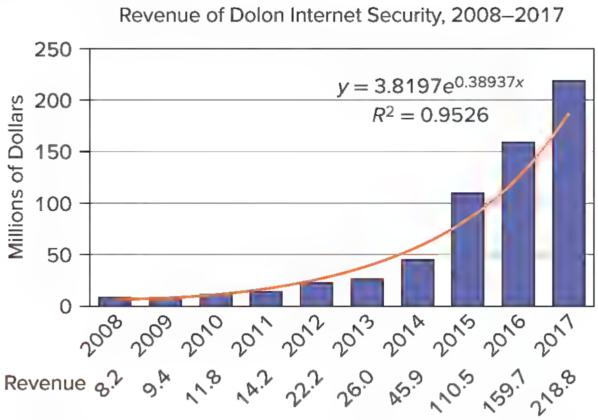
Illustration: Exponential Trend

Spending on Internet security in the United States has shown explosive growth. For example, Figure 14.11 shows revenue growth for one web security company. Clearly, a linear trend (constant *dollar* growth) would be inadequate. It is more reasonable to assume a constant

*In a sense, the linear model ($y_t = a + bt$) and the exponential model ($y_t = ae^{bt}$) are equally simple because they are two-parameter models, and a log-transformed exponential model $\ln(y_t) = \ln(a) + bt$ is actually linear.

Figure 14.11

Excel's Exponential Trend DolonCorp



percent rate of growth and fit an exponential model. Excel's fitted exponential trend is $y_t = 3.8197e^{0.3894t}$. The value of b in the exponential model $y_t = ae^{bt}$ is the continuously compounded growth rate, so we can say that Dolon's revenue is growing at an astonishing rate of 38.94 percent per year. A negative value of b in the equation $y_t = ae^{bt}$ would indicate *decline* instead of growth. The intercept a is the "starting point" in period $t = 0$. For example, $y_0 = 3.8197\, e^{0.3894(0)} = 3.8197$.

Exponential Trend Calculations

Table 14.5 shows the worksheet for the required sums. Calculations of the exponential trend are done by using a transformed variable $z_t = \ln(y_t)$, instead of y_t, to produce a linear equation so that we can use the least squares formulas.

$$\textit{Slope:}\quad b = \frac{\sum_{t=1}^{n}(t-\bar{t})(z_t-\bar{z})}{\sum_{t=1}^{n}(t-\bar{t})^2} = \frac{32.12329}{82.5} = 0.3893732$$

$$\textit{Intercept:}\ a = \bar{z} - b\bar{t} = 3.481731 - (.3893732)(5.5) = 1.340178$$

When the least squares calculations are completed, we must transform the intercept back to the original units by exponentiation to get the correct intercept $a = e^{1.340178} = 3.8197$. In final form, the fitted trend equation is

$$y_t = ae^{bt} = 3.8197e^{0.3894t}$$

Table 14.5

Least Squares Sums for the Exponential Model DolonCorp

Year	t	y_t	$z_t = \ln(y_t)$	$t-\bar{t}$	$z_t-\bar{z}$	$(t-\bar{t})^2$	$(t-\bar{t})(z_t-\bar{z})$
2008	1	8.2	2.10413	−4.5	−1.37760	20.25	6.19919
2009	2	9.4	2.24071	−3.5	−1.24102	12.25	4.34357
2010	3	11.8	2.46810	−2.5	−1.01363	6.25	2.53408
2011	4	14.2	2.65324	−1.5	−0.82849	2.25	1.24273
2012	5	22.2	3.10009	−0.5	−0.38164	0.25	0.19082
2013	6	26.0	3.25810	0.5	−0.22363	0.25	−0.11182
2014	7	45.9	3.82647	1.5	0.34473	2.25	0.51710
2015	8	110.5	4.70502	2.5	1.22328	6.25	3.05821
2016	9	159.7	5.07330	3.5	1.59157	12.25	5.57048
2017	10	218.8	5.38816	4.5	1.90643	20.25	8.57892
Sum	55	626.7	34.81731	0.0	0.00000	82.5	32.12329
Mean	5.5	62.67	3.481731				

Forecasting an Exponential Trend

We can make a forecast of debit card usage for any future year by using the fitted model*:

For 2018 ($t = 11$): $y_{11} = 3.8197e^{0.38937(11)} = 276.8$

For 2019 ($t = 12$): $y_{12} = 3.8197e^{0.38937(12)} = 408.5$

For 2020 ($t = 13$): $y_{13} = 3.8197e^{0.38937(13)} = 603.0$

Can Dolon's revenue actually continue to grow at a rate of 38.937 percent? It seems unlikely. Typically, when a new product is introduced, its growth rate at first is very strong but eventually slows down as the market becomes saturated and/or as competitors arise. While the 2018 and 2019 projections would be reasonable, the 2020 forecast ignores the impact of the COVID-19 crisis. This reminds us that trend forecasts are useful only if the past trend can reasonably be assumed to provide a reliable guide to the near future.

Exponential Trend: Calculating R^2

As shown in Table 14.6, we calculate R^2 the same way as for the linear trend, except that we replace the dependent variable y_t with $z_t = \ln(y_t)$ and the fitted value with $\hat{z}_t = 1.340178 + .389373t$. This is necessary because Excel's trend-fitting calculations are done in logarithms:

$$\textit{Coefficient of determination:} \quad R^2 = 1 - \frac{\sum_{t=1}^{n}(z_t - \hat{z}_t)^2}{\sum_{t=1}^{n}(z_t - \bar{z})^2} = 1 - \frac{0.62228}{13.13021} = 0.9526$$

In this example, the exponential trend gives a very good fit ($R^2 = 0.9526$) to the past data. Although a high R^2 does not guarantee good forecasts, demand for Internet security protection is expected to grow, so Dolon's high growth rate could continue if the firm is able to manage its expansion.

Table 14.6 **Sums for R^2 Calculations in Exponential Model** DolonCorp

t	$z_t = \ln(y_t)$	$\hat{z}_t = 1.340178 + .389373t$	$z_t - \hat{z}_t$	$(z_t - \hat{z}_t)^2$	$(z_t - \bar{z})^2$
1	2.10413	1.72955	0.37458	0.14031	1.89777
2	2.24071	2.11892	0.12178	0.01483	1.54013
3	2.46810	2.50830	−0.04020	0.00162	1.02745
4	2.65324	2.89767	−0.24443	0.05975	0.68639
5	3.10009	3.28704	−0.18695	0.03495	0.14565
6	3.25810	3.67642	−0.41832	0.17499	0.05001
7	3.82647	4.06579	−0.23933	0.05728	0.11884
8	4.70502	4.45516	0.24985	0.06243	1.49643
9	5.07330	4.84454	0.22876	0.05233	2.53308
10	5.38816	5.23391	0.15425	0.02379	3.63446
Sum	34.81731	34.81731	0	0.62228	13.13021
Mean	3.48173				

Knowing only y_1 and y_t (the starting and ending values) you can estimate the compound growth rate b using this formula:

$$b = [\ln(y_t) - \ln(y_1)]/(t - 1) \qquad \textbf{(14.4)}$$

*Excel uses the exponential formula $y_t = ae^{bt}$, in which the coefficient b is the *continuously compounded* growth rate. Minitab uses $y_t = y_0(1 + r)^t$, which you may recognize as the formula for compound interest. Although the formulas appear different, they give identical forecasts. To convert Minitab's fitted equation to Excel's, set $a = y_0$ and $b = \ln(1 + r)$. To convert Excel's fitted equation to Minitab's, set $y_0 = a$ and $r = e^b - 1$.

We can apply this formula to Dolon's revenue using Excel's natural log function:

$$b = (\text{LN}(218.8) - \text{LN}(8.2))/(10 - 9) = 0.3649$$

This formula is useful (e.g., for comparing investments) when you only know where you started and where you are now.

Quadratic Trend Model

The **quadratic trend** model has the form $y_t = a + bt + ct^2$. The t^2 term allows a nonlinear shape. It is useful for a time series that has a turning point or that is not captured by the exponential model. If $c = 0$, the quadratic model $y_t = a + bt + ct^2$ becomes a linear model because the term ct^2 drops out of the equation (i.e., the linear model is a special case of the quadratic model). Fitting a quadratic model is a way of checking for nonlinearity. If the coefficient c does not differ significantly from zero (and if the quadratic R^2 is about the same as a linear model), then the linear model would suffice. Depending on the values of b and c, the quadratic model can assume any of four shapes, as shown in Figure 14.12.

Figure 14.12

Four Quadratic Trend Models

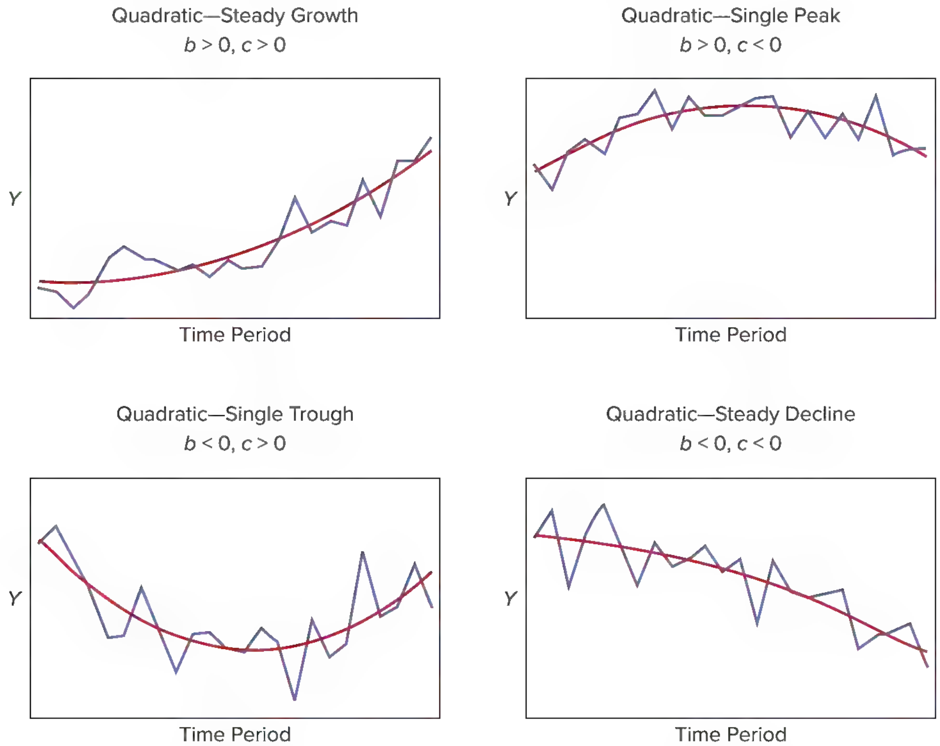


Illustration: Quadratic Trend

The number of hospital beds (Table 14.7) in the United States declined during the late 1990s, showed signs of leveling out, and then declined again. What trend would we choose if the objective is to make a realistic one-year forecast?

Table 14.7

U.S. Hospital Beds (thousands), 1995–2004

HospitalBeds

Source: *Statistical Abstract of the United States, 2007,* p. 114.

Year	Beds	Year	Beds
1995	1,081	2000	984
1996	1,062	2001	987
1997	1,035	2002	976
1998	1,013	2003	965
1999	994	2004	956

Figure 14.13

Two Trend Models for U.S. Hospital Beds, 1995–2004 **HospitalBeds**

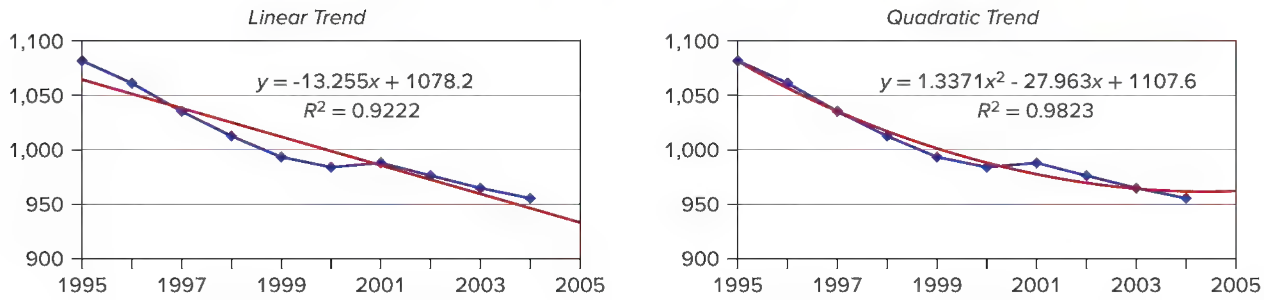


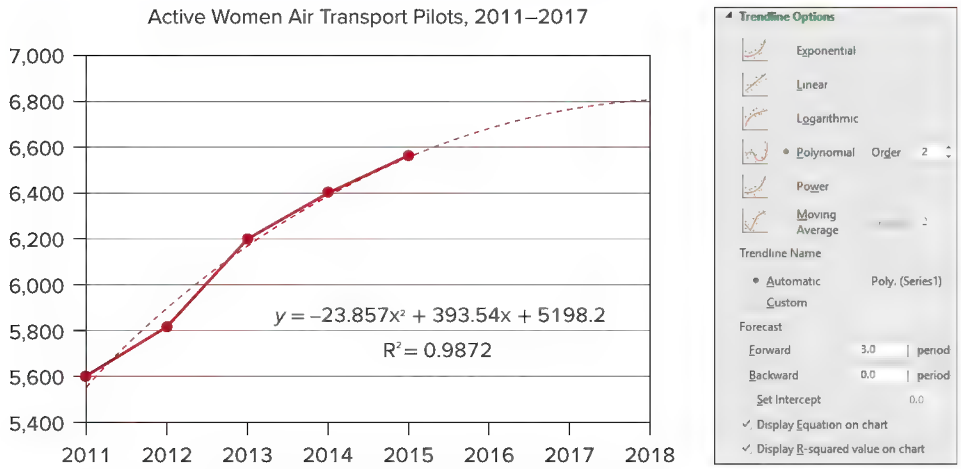
Figure 14.13 shows one-year projections using the linear and quadratic models. Many observers would think that the quadratic model offers a more believable prediction because the quadratic model is able to capture the slight curvature in the data pattern. But this gain in credibility must be weighed against the added complexity of the quadratic model. It appears that the forecasts would turn upward if projected more than one year ahead. We should be especially skeptical of any polynomial model that is projected more than one or two periods into the future.

Using Excel for Trend Fitting

Plot the data, right-click on the data, and choose a trend. Figure 14.14 shows Excel's menu of six trend options. The menu includes a sketch of each trend type. Click the Options tab if you want to display the R^2 and fitted equation on the graph, or if you want to plot forecasts (trend extrapolations) on the graph. The quadratic model is a **polynomial model** of order 2. Despite Excel's many choices, some patterns cannot be captured by any of the common trend models. For women air transport pilots, the fitted quadratic (polynomial) regression predicts continued growth, but at a slowing rate. By default, Excel reports four- or five-decimal accuracy. However, you can click on Excel's fitted trend equation, choose Format Data Labels, choose Number, and set the number of decimal places you want to see.

Figure 14.14

Excel's Trend-Fitting Menus **WomenPilots**



Source: Microsoft Excel.

Principle of Occam's Razor

Given two *sufficient* explanations, we prefer the simpler one.

William of Occam (1285–1347)

Trend-Fitting Criteria

It is so easy to fit a trend in Excel that it is tempting to "shop around" for the best fit. Forecasters prefer the simplest trend model that adequately matches the trend. Simple models are easier to interpret and explain to others. However, that is *not* to say that a simpler model is always preferred. Occam's Razor is merely a "tie-breaker" when we have two *equally* good models. Criteria for selecting a trend model for forecasting include

Criterion	*Ask Yourself*
• Occam's Razor	Would a simpler model suffice?
• Overall fit	How does the trend fit the past data?
• Believability	Does the extrapolated trend "look right"?
• Fit to recent data	Does the fitted trend match the last few data points?

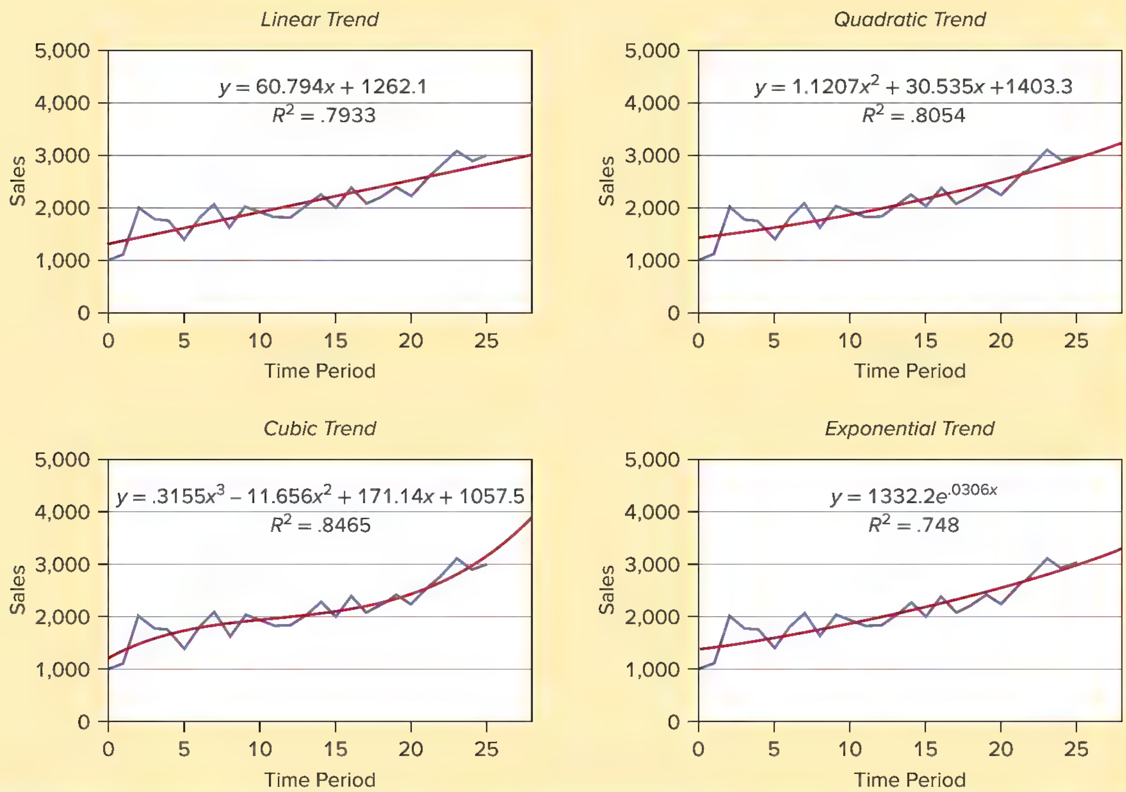
EXAMPLE 14.1

Comparing Trends

You can usually increase the R^2 by choosing a more complex model. But if you are making a *forecast,* this is not the only relevant issue because R^2 measures the fit to the *past* data. Figure 14.15 shows four fitted trends using the same data, with three-period

Figure 14.15

Four Fitted Trends Using the Same Data



forecasts. For this data set, the linear model may be inadequate because its fit to recent periods is marginal (we prefer the simplest model *only if* it "does the job"). The cubic trend yields the highest R^2, but the fitted equation is nonintuitive and would be hard to explain or defend. Also, its forecasts appear to be increasing too rapidly. In this example, the exponential model has the lowest R^2, yet it matches the recent data fairly well and its forecasts appear credible when projected a few periods ahead.

Any trend model's forecasts become less reliable as they are extrapolated farther into the future. The quadratic trend, the simplest of Excel's polynomial models, is sometimes acceptable for short-term forecasting. However, forecasters avoid higher-order polynomial models (cubic and higher) not only because they are complex but also because they can give bizarre forecasts when extrapolated more than one period ahead. Table 14.8 compares the features of common trend models.

Table 14.8

Comparison of Three Trend Models

Model	Pro	Con
Linear	1. Simple, familiar to everyone. 2. May suffice for short-run data.	1. Assumes constant slope. 2. Cannot capture nonlinear change.
Exponential	1. Familiar to financial analysts. 2. Shows compound growth rate.	1. Some managers are unfamiliar with e^x. 2. Data values must be positive.
Quadratic	1. Useful for data with a turning point. 2. Useful test for nonlinearity.	1. Complex and lacks intuitive interpretation. 2. Can give untrustworthy forecasts if extrapolated too far.

Analytics in Action

Trend? Or Bubble?

A price "bubble" is a rapid increase in the price of an asset well above its typical market value, which creates instability in the market. Rising prices may inspire investors to seek credit to finance their increased purchases of the asset. Rising prices may also convince banks and capital markets to accommodate these requests by issuing credit. However, a bubble can collapse if creditors become alarmed and stop allowing the asset to be used as collateral. A bubble is especially fragile when financing is highly leveraged. The U.S. housing bubble peaked in 2006 and reached new lows by 2012. Are similar bubbles occurring now? A recent Forbes article was titled "The S&P 500 Bubble Is Coming: What Now?" When does a surge in price of cryptocurrency (e.g., Bitcoin) become a bubble? What about recent surging prices of rare metals (e.g., rhodium)? How can we tell when an asset price increase is a bubble?

Some market observers have suggested that prices more than *two standard deviations* above the longer trend may signal a bubble. You know how to fit trends and construct

95 percent confidence intervals, so you can do your own bubble hunting. Of course, everyone else can do the same. Nobel prize–winning economist Robert Shiller of Yale University has published extensively on using analytics to develop models of asset valuation. His models are far from simple. One of Shiller's insights is that in efficient markets everyone has access to the same information and analytical tools. Be forewarned—there is no simple formula for investment success. And it's psychology—not just economics.

Source: Robert J. Shiller, *Narrative Economics: How Stories Go Viral and Drive Major Economic Events.* (Princeton, NJ: Princeton University Press, 2019).

Section Exercises

connect

14.1 In 2009, US Airways Flight 1549 made a successful emergency landing in the Hudson River, after striking birds shortly after takeoff. Are bird strikes an increasing threat to planes? (a) Make an Excel graph of the data on bird strikes. (b) Discuss the underlying causes that might explain the trend. (c) Fit three trends (linear, quadratic, and exponential) to the time series. (d) Use *each* of the three fitted trend equations to make numerical forecasts for the next three years. How much difference does the choice of model make? Which forecasts do you trust the most, and why? **BirdStrikes**

Number of Reported Bird Strikes to Civil Aircraft in U.S., 2008–2018 **BirdStrikes**

Year	Strikes	Year	Strikes	Year	Strikes
2008	7,213	2012	10,918	2016	13,454
2009	8,950	2013	11,417	2017	14,664
2010	9,905	2014	13,694	2018	16,020
2011	10,119	2015	13,808		

Source: http://wildlife-mitigation.tc.faa.gov.

14.2 (a) Make an Excel graph of the data on usage of renewable energy in the United States. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Fit three trends (linear, quadratic, exponential) to the time series. (d) Use *each* of the three fitted trend equations to make numerical forecasts for the next three years. How similar are the three models' forecasts? **Renew**

U.S. Usage of Renewable Energy (quad BTU), 2011–2018

Year	Usage	Year	Usage
2011	9.20	2015	9.72
2012	8.85	2016	10.37
2013	9.45	2017	11.18
2014	9.74	2018	11.52

Source: https://usafacts.org/data/.

14.3 (a) Make an Excel line graph of the data on employee work stoppages. (b) Discuss the underlying causes that might explain the trend or pattern. (c) Fit three trends (linear, exponential, quadratic). (d) Which trend model is best, and why? If none is satisfactory, explain. (e) Would you trust a trend forecast for 2020? Explain. **Strikers**

U.S. Workers Involved in Work Stoppages, 2000–2019 (thousands)

Year	Strikers	Year	Strikers	Year	Strikers	Year	Strikers
2000	397	2005	102	2010	45	2015	49
2001	102	2006	77	2011	113	2016	102

U.S. Workers Involved in Work Stoppages, 2000–2019 (thousands)

Year	Strikers	Year	Strikers	Year	Strikers	Year	Strikers
2002	47	2007	193	2012	150	2017	25
2003	131	2008	83	2013	55	2018	485
2004	232	2009	13	2014	34	2019	466

Source: http://data.bls.gov.

14.4 You want to invest $1,000. Which growth curve would yield the largest principal 5 years from now? 10 years? 20 years? Explain. *Hint:* Show all the forecasts.

a. $y_t = 1000e^{0.039t}$
b. $y_t = 1000 + 45t$
c. $y_t = 1000 + 11t + 3t^2$

14.5 For each of the following fitted trends, make a prediction for period $t = 15$:

a. $y_t = 926e^{-.026t}$
b. $y_t = 2{,}217 - 8t$
c. $y_t = 447 - 29t + 7t^2$

Mini Case 14.1

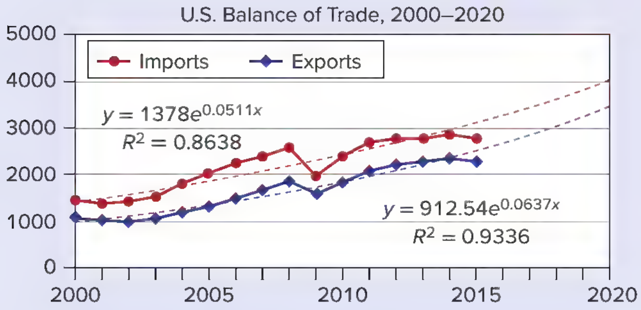
U.S. Trade Deficit

The imbalance between imports and exports has been a vexing policy problem for U.S. policymakers for decades. The last time the United States had a trade surplus was in 1975, partly due to reduced dependency on foreign oil through conservation measures enacted after the oil crisis (shortages and gas lines) in the early 1970s. However, the trade deficit has become more acute over time due partly to continued oil imports and, more recently, to availability of cheaper goods from China and other emerging economies.

Prior to the recent recession, imports had been growing faster than exports. Yet over the past decade, the fitted trend equations show that exports have grown at a slightly higher compound annual rate of 6.37 percent, compared with 5.11 percent for imports (see Figure 14.16). Possible reasons would include reduced industrial production due to the recession, improved vehicle fuel economy, rising import prices, and the impact of exchange rates.

Figure 14.16

U.S. Trade, 2000–2020 TradeDeficit



Perhaps the United States can achieve trade balance–but only far in the future, given the tiny difference in growth rates. Further, the assumption of *ceteris paribus* may not hold. Much depends on U.S. trade treaties, global challenges (e.g., climate change), how the United States and other nations handle their internal finances, and international conflicts. The disruptive COVID-19 crisis of 2020 illustrates the difficulty of making economic policy based on trend projections.

Year	t	Exports (projection)	Imports (projection)
2016	$t = 17$	y_{17} = 912.54*EXP(0.0637*17) = 2,695	y_{17} = 1378*EXP(0.0511*17) = 3,285
2017	$t = 18$	y_{18} = 912.54*EXP(0.0637*18) = 2,872	y_{18} = 1378*EXP(0.0511*18) = 3,457
2018	$t = 19$	y_{19} = 912.54*EXP(0.0637*19) = 3,061	y_{19} = 1378*EXP(0.0511*19) = 3,638
2019	$t = 20$	y_{20} = 912.54*EXP(0.0637*20) = 3,762	y_{20} = 1378*EXP(0.0511*20) = 3,829
2020	$t = 21$	y_{21} = 912.54*EXP(0.0637*21) = 3,477	y_{21} = 1378*EXP(0.0511*21) = 4,030

Forecasts are less a way of predicting the future than of showing where we are heading *if* nothing changes. A paradox of forecasting is that, as soon as decision makers see the implications of a distasteful forecast, they may try to take steps to ensure that the forecast is wrong!

14.3 ASSESSING FIT

LO 14-4

Know the definitions of common fit measures.

Five Measures of Fit

In time-series analysis, you are likely to encounter several different measures of "fit" that show how well the estimated trend model matches the observed time series. "Fit" refers to historical data, and you should bear in mind that a good fit is no guarantee of good forecasts–the usual goal. Five common measures of fit are shown in Table 14.9.

Table 14.9

Five Measures of Fit

Statistic	Description	Pro	Con
(14.5) $R^2 = 1 - \dfrac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$	**Coefficient of determination (R^2)**	1. Unit-free measure. 2. Very common.	1. Often interpreted incorrectly (e.g., "percent of correct predictions").
(14.6) $MAPE = \dfrac{100}{n}\sum_{t=1}^{n}\dfrac{\lvert y_t - \hat{y}_t\rvert}{y_t}$	**Mean absolute percent error (*MAPE*)**	1. Unit-free measure (%). 2. Intuitive meaning.	1. Requires $y_t > 0$. 2. Lacks nice math properties.
(14.7) $MAD = \dfrac{1}{n}\sum_{t=1}^{n}\lvert y_t - \hat{y}_t\rvert$	**Mean absolute deviation (*MAD*)**	1. Intuitive meaning. 2. Same units as y_t.	1. Not unit-free. 2. Lacks nice math properties.
(14.8) $MSD = \dfrac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$	**Mean squared deviation (*MSD*)**	1. Nice math properties. 2. Penalizes big errors more.	1. Nonintuitive meaning. 2. Rarely reported.
(14.9) $SE = \sqrt{\sum_{t=1}^{n}\dfrac{(y_t - \hat{y}_t)^2}{n-2}}$	**Standard error (*SE*)**	1. Same units as y_t. 2. For confidence intervals.	1. Nonintuitive meaning.

EXAMPLE 14.2

Fire Losses

Figure 14.17 shows an Excel graph with fitted linear trend and three-year forecasts for fire loss claims paid to homeowners (in millions of dollars) by an insurance company. Table 14.10 shows the calculations for these statistics of fit. Because the residuals $y_t - \hat{y}_t$ sum to zero, we see why it's necessary to sum either their absolute values or their squares to obtain a measure of fit. *MAPE, MAD, MSD,* and *SE* would be zero if the trend provided a perfect fit to the time series.

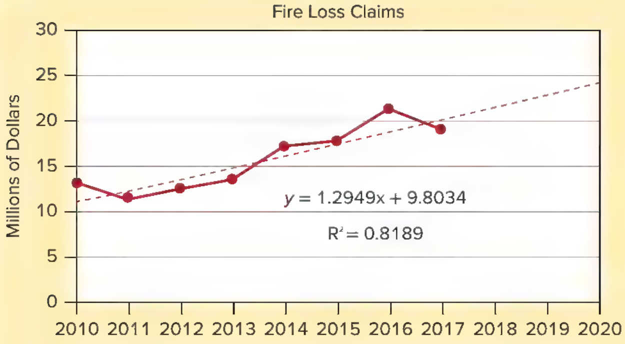


Figure 14.17

Fire Loss Claims Paid—Linear Model
FireLosses

Using the sums in Table 14.10, we can apply the formulas for each fit statistic:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t} = \frac{100}{8}(0.6590) = 8.24\%$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| = \frac{1}{8}(9.9142) = 1.239$$

$$MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 = \frac{1}{8}(15.5683) = 1.946$$

$$SE = \sqrt{\sum_{t=1}^{n}\frac{(y_t - \hat{y}_t)^2}{n-2}} = \sqrt{\frac{15.5683}{8-2}} = 1.611$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} = 1 - \frac{15.5683}{85.9878} = .8189$$

The R^2 statistic says that a linear trend alone can "explain" about 82 percent of the variation in claims paid. *MAPE* says that our fitted trend has a mean absolute error of 8.24 percent. *MAD* says that the average error is 1.239 million dollars (ignoring the sign). *MSD* lacks a simple interpretation. These fit statistics are most useful in comparing different trend models for the same data. All the statistics (especially the *MSD*) are affected by the unusual residual in 2016, when fire losses greatly exceeded the trend. The standard error is useful if we want to make a prediction interval for a forecast, using formula 14.9. It is the same formula you saw in Chapter 12. This formula widens the confidence interval when the time index t is far from its historic mean.

$$\hat{y}_t \pm t_{n-2}SE\sqrt{1 + \frac{1}{n} + \frac{(t-\bar{t})^2}{\sum_{t=1}^{n}(t-\bar{t})^2}} \quad \text{(prediction interval for future } y_t) \qquad \textbf{(14.10)}$$

Table 14.10

SUMS for *MAD, MAPE, MSD,* and Standard Error **FireLosses**

Period	Year	y_t	$\hat{y}_t = 9.8034 + 1.2949t$		$y_t - \hat{y}_t$	$\lvert y_t - \hat{y}_t\rvert$	$\lvert y_t - \hat{y}_t\rvert / y_t$	$(y_t - \hat{y}_t)^2$
1	2010	12.940	11.0983		1.8417	1.8417	0.1423	3.3919
2	2011	11.510	12.3932		−0.8832	0.8832	0.0767	0.7800
3	2012	12.428	13.6881		−1.2601	1.2601	0.1014	1.5879
4	2013	13.457	14.9830		−1.5260	1.5260	0.1134	2.3287
5	2014	17.118	16.2779		0.8401	0.8401	0.0491	0.7058
6	2015	17.586	17.5728		0.0132	0.0132	0.0008	0.0002
7	2016	21.129	18.8677		2.2613	2.2613	0.1070	5.1135
8	2017	18.874	20.1626		−1.2886	1.2886	0.0683	1.6605
				Sum	0.000	9.9142	0.6590	15.5683
				Mean	0.000	1.2393	0.0824	1.9460

14.4 Moving Averages

LO 14-5

Interpret a moving average and use Excel to create it.

Trendless or Erratic Data

What if the time series $y_1, y_2, \ldots, y_n$ is erratic or has no consistent trend? In such cases, there may be little point in fitting a trend, and if the mean is changing over time, we cannot just "take the average" over the entire data set. Instead, a conservative approach is to calculate a **moving average**. There are two main types of moving averages: trailing or centered. We will illustrate each.

Trailing Moving Average (*TMA*)

The simplest kind of moving average is the **trailing moving average** (*TMA*) over the last m periods.

(14.11) $$\hat{y}_t = \frac{y_t + y_{t-1} + \ldots + y_{t-m+1}}{m}$$ (trailing moving average over m periods)

The *TMA* smooths the past fluctuations in the time series, helping us see the pattern more clearly. The choice of m depends on the situation. A larger m yields a "smoother" *TMA* but requires more data. The value of $\hat{y}_t$ also may be used as a forecast for period $t + 1$. Beyond the range of the observed data $y_1, y_2, \ldots, y_n$ there is no way to update the moving average, so it is best regarded as a *one-period-ahead forecast.*

EXAMPLE 14.3

Fuel Economy

Many drivers keep track of their fuel economy. For a given vehicle, there is likely to be little trend over time, but there is always random fluctuation. Also, current driving conditions (e.g., snow, hot weather, road trips) could temporarily affect mileage over several consecutive time periods. In this situation, a moving average might be considered. Table 14.11 shows Andrew's fuel economy data set. Column five shows a three-period *TMA*. For example, for period 6 (yellow-shaded cells), the *TMA* is

$$\hat{y}_6 = \frac{24.392 + 21.458 + 24.128}{3} = 23.326$$

It is easiest to appreciate the moving average's "smoothing" of the data when it is displayed on a graph, as in Figure 14.18. It is clear that Andrew's mean is around 23 mpg, though the moving average fluctuates over a range of approximately ± 2 mpg.

Figure 14.18

Three-Period Moving Average of MPG

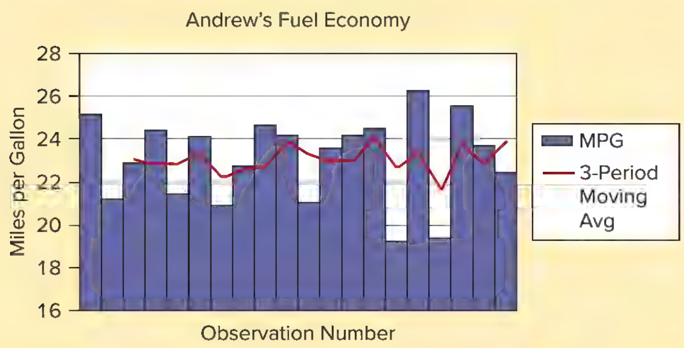


Table 14.11 Andrew's Miles per Gallon ($n = 20$) AndrewsMPG

Obs	Date	Miles Driven	Gallons	MPG	TMA	CMA
1	5-Jan	285	11.324	25.168		
2	7-Jan	185	8.731	21.189		23.074
3	11-Jan	250	10.934	22.864	23.074	22.815
4	15-Jan	296	12.135	24.392	22.815	22.905
5	19-Jan	232	10.812	21.458	22.905	23.326
6	25-Jan	301	12.475	24.128	23.326	22.158
7	30-Jan	285	13.645	20.887	22.158	22.581
8	3-Feb	263	11.572	22.727	22.581	22.747
9	7-Feb	250	10.152	24.626	22.747	23.856
10	14-Feb	307	12.678	24.215	23.856	23.283
11	22-Feb	242	11.520	21.007	23.283	22.942
12	29-Feb	288	12.201	23.605	22.942	22.937
13	5-Mar	285	11.778	24.198	22.937	24.103
14	8-Mar	313	12.773	24.505	24.103	22.638
15	13-Mar	283	14.732	19.210	22.638	23.330
16	18-Mar	318	12.103	26.274	23.330	21.620
17	22-Mar	195	10.064	19.376	21.620	23.746
18	28-Mar	320	12.506	25.588	23.746	22.904
19	2-Apr	270	11.369	23.749	22.904	23.910
20	12-Apr	259	11.566	22.393	23.910	

Source: Data were collected by statistics student Andrew Fincher for his 11-year-old car.

Example: *TMA*

Example: *CMA*

Centered Moving Average (*CMA*)

Another moving average is the **centered moving average (*CMA*)**. Formula 14.12 shows a *CMA* for $m = 3$ periods. The formula looks both forward *and* backward in time, to express the current "forecast" as the mean of the current observation *and* observations on either side of the current data.

(14.12) $$\hat{y}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3} \quad \text{(centered moving average over } m \text{ periods)}$$

This is not really a forecast at all, but merely a way of smoothing the data. In Table 14.11, column seven shows the *CMA* for Andrew's MPG data. For example, for period 14 (blue-shaded cells), the *CMA* is

$$\hat{y}_t = \frac{24.198 + 24.505 + 19.210}{3} = 22.638$$

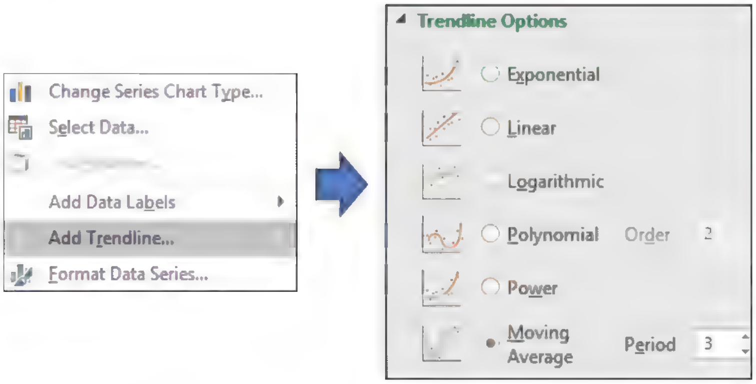
When n is odd ($m = 3, 5$, etc.), the *CMA* is easy to calculate. When m is even, the formula is more complex because the mean of an even number of data points would lie *between* two data points and would not be correctly centered. Instead, we take a double moving average (yipe!) to get the resulting *CMA* centered properly. For example, for $m = 4$, we would average y_{t-2} through y_{t+1}, then average y_{t-1} through y_{t+2}, and finally average the two averages! You need not worry about this formula for now. It will be illustrated shortly in the context of seasonal data.

Using Excel for a TMA

Excel offers a *TMA* in its Add Trendline option when you click on a time-series line graph or bar chart. Its menus are displayed in Figure 14.19. The *TMA* is a conservative choice whenever you doubt that one of Excel's five other trend models (linear, logarithmic, polynomial, power, exponential) would be appropriate. However, Excel does *not* give you the option of making any forecasts with its moving average model.

Figure 14.19

Excel's Moving Average Menus



Source: Microsoft Excel

Section Exercise

connect

14.6 (a) Make an Excel line graph of the exchange rate data (only first 3 and last 3 days are shown). Describe the pattern. (b) Click on the data and choose Add Trendline > Moving Average. Describe the effect of increasing m (e.g., $m = 2, 4, 6$, etc.). Include a copy of each graph with your answer. (c) Discuss how this moving average might help a currency speculator. **DollarEuro**

Daily Dollar/Euro Exchange Rate ($n = 61$ days)

Day	1	2	3	. . .	59	60	61
Date	11/1/19	11/4/19	11/5/19	. . .	1/29/20	1/30/20	1/31/20
Rate	1.1169	1.1144	1.1070	. . .	1.1004	1.1032	1.1082

Source: www.federalreserve.gov.

14.5 EXPONENTIAL SMOOTHING

LO 14-6

Use exponential smoothing to forecast trendless data.

Forecast Updating

The **exponential smoothing** model is a special kind of moving average. It is used for ongoing one-period-ahead forecasting for data that have up-and-down movements but no consistent trend. For example, a retail outlet may place orders for thousands of different stock-keeping units (SKUs) each week, so as to maintain its inventory of each item at the desired level (to avoid emergency calls to warehouses or suppliers). For such forecasts, many firms choose exponential smoothing, a simple forecasting model with only two inputs and one constant. The updating formula for the forecasts is

(14.13) $$F_{t+1} = \alpha y_t + (1 - \alpha) F_t \quad \text{(Smoothing update)}$$

where

F_{t+1} = the forecast for the next period

α = the "smoothing constant" ($0 \le \alpha \le 1$)

y_t = the actual data value in period t

F_t = the previous forecast for period t

Smoothing Constant (α)

The next forecast F_{t+1} is a weighted average of y_t (the current data) and F_t (the previous forecast). The value of α, called the **smoothing constant**, is the weight given to the latest data. A small value of α would give low weight to the most recent observation and heavy weight $1 - \alpha$ to the previous forecast (a "heavily smoothed" series). The larger the value of α, the more quickly the forecasts adapt to recent data. For example,

If $\alpha = .05$, then $F_{t+1} = .05y_t + .95F_t$ (heavy smoothing, slow adaptation)
If $\alpha = .20$, then $F_{t+1} = .20y_t + .80F_t$ (moderate smoothing, moderate adaptation)
If $\alpha = .50$, then $F_{t+1} = .50y_t + .50F_t$ (little smoothing, quick adaptation)

Choosing the Value of α

If $\alpha = 1$, there is no smoothing at all, and the forecast for next period is the same as the latest data point, which basically defeats the purpose of exponential smoothing. Minitab uses $\alpha = .20$ (i.e., moderate smoothing) as its default, which is a fairly common choice of α. The fit of the forecasts to the data will change as you try different values of α. Most computer packages can, as an option, solve for the "best" α using a criterion such as minimum *SSE.*

Over time, earlier data values have less effect on the exponential smoothing forecasts than more recent y-values. To see this, we can replace F_t in formula 14.12 with the prior forecast F_{t-1}, and repeat this type of substitution indefinitely to obtain this result:

$$F_{t+1} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \cdots \qquad \textbf{(14.14)}$$

We see that the next forecast F_{t+1} depends on *all* the prior data (y_{t-1}, y_{t-2}, etc.). However, as long as $\alpha < 1$, as we go farther into the past, each prior data value has less and less impact on the current forecast.

Initializing the Process

From formula 14.12, we see that F_{t+1} depends on F_t, which in turn depends on F_{t-1}, and so on, all the way back to F_1. But where do we get F_1 (the initial forecast)? There are many ways to initialize the forecasting process. For example, Excel simply sets the initial forecast equal to the first actual data value:

Method *A*
Set $F_1 = y_1$ (use the first data value)

This method has the advantage of simplicity, but if y_1 happens to be unusual, it could take a few iterations for the forecasts to stabilize. Another approach is to set the initial forecast equal to the average of the first several observed data values. For example, Minitab uses the first six data values:

Method *B*

Set $F_1 = \dfrac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{n}$ (average of first 6 data values)

This method tends to iron out the effects of unusual y-values, but it consumes more data and is still vulnerable to unusual y-values.

EXAMPLE 14.4

Weekly Sales Data

Table 14.12 shows weekly sales of deck sealer (a paint product sold in gallon containers) at a large do-it-yourself warehouse-style retailer. For exponential smoothing forecasts, the company uses $\alpha = .10$. Its choice of α is based on experience. Because α is fairly small, it will provide strong smoothing. The last two columns compare the two methods of initializing the forecasts. Unusually high sales in week 5 have a strong effect on method *B*'s starting point. At first, the difference in forecasts is striking, but over time the methods converge.

Using Method A:

$$F_2 = \alpha y_1 + (1-\alpha)F_1 = (.10)(106) + (.90)(106) = 106$$
$$F_3 = \alpha y_2 + (1-\alpha)F_2 = (.10)(110) + (.90)(106) = 106.4$$
$$F_4 = \alpha y_3 + (1-\alpha)F_3 = (.10)(108) + (.90)(106.4) = 106.56$$
$$\vdots$$
$$F_{19} = \alpha y_{18} + (1-\alpha)F_{18} = (.10)(120) + (.90)(130.908) = 129.82$$

Table 14.12 Deck Sealer Sales: Exponential Smoothing (n = 18 weeks)
DeckSealer

Week	Actual Sales	Method A: $F_1 = y_1$	Method B: F_1 = Average (1st 6)
1	106	106.000	127.833
2	110	106.000	125.650
3	108	106.400	124.085
4	97	106.560	122.477
5	210	105.604	119.929
6	136	116.044	128.936
7	128	118.039	129.642
8	134	119.035	129.478
9	107	120.532	129.930
10	123	119.179	127.637
11	139	119.561	127.174
12	140	121.505	128.356
13	144	123.354	129.521
14	94	125.419	130.969
15	108	122.277	127.272
16	168	120.849	125.344
17	179	125.564	129.610
18	120	130.908	134.549

Smoothed forecasts using $\alpha = .10$.

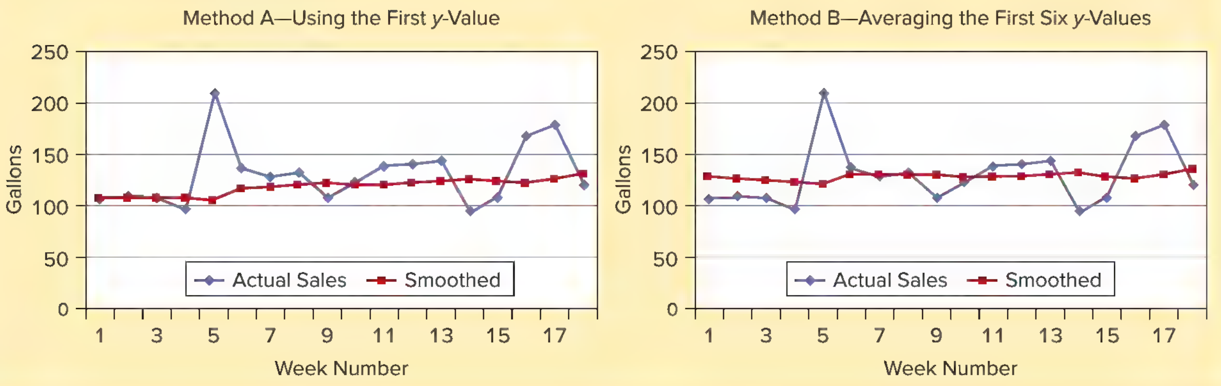
Using Method B:

$$F_2 = \alpha y_1 + (1-\alpha)F_1 = (.10)(106) + (.90)(127.833) = 125.650$$
$$F_3 = \alpha y_2 + (1-\alpha)F_2 = (.10)(110) + (.90)(125.650) = 124.085$$
$$F_4 = \alpha y_3 + (1-\alpha)F_3 = (.10)(108) + (.90)(124.085) = 122.477$$
$$\vdots$$
$$F_{19} = \alpha y_{18} + (1-\alpha)F_{18} = (.10)(120) + (.90)(134.549) = 133.094$$

Despite their different starting points, the forecasts for period 19 do not differ greatly. Rounding to the next higher integer, for week 19, the firm would order 130 gallons (using method *A*) or 134 gallons (using method *B*). Figure 14.20 shows the similarity in *patterns* of the forecasts, although the *level* of forecasts is always higher in method *B* because of its higher initial value. This demonstrates that the choice of starting values *does* affect the forecasts.

Figure 14.20

Initializing Methods Compared



Using Excel

Excel has an exponential smoothing option. It is found in the Data Analysis menu. Instead of the smoothing constant α, Excel asks for a *damping factor,* which is equal to 1 − α. Excel uses method *A* (setting $F_1 = y_1$) to initialize its forecasts. Figure 14.22 shows Excel's exponential smoothing dialogue box and its output chart of actual and forecast values. Excel's chart is difficult to read, so you may wish to make your own "improved" line chart, like the one shown in Figure 14.21. Excel makes no *future* forecasts past period $t = 18$, but you can do it yourself (see the period $t = 19$ forecast calculations in Example 14.4). Exponential smoothing forecasts can't be updated beyond one period ahead because there are no more actual y_t values to plug into the updating formula.

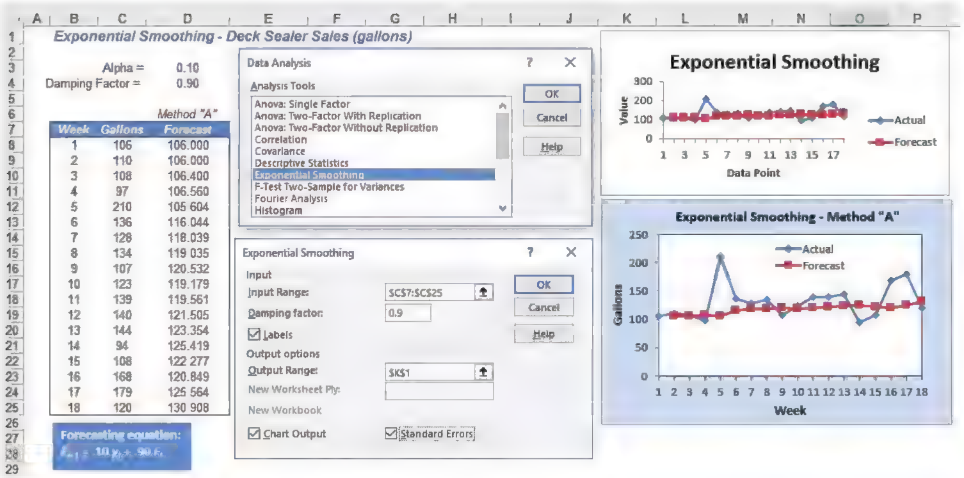


Figure 14.21

Excel's Exponential Smoothing

Source: Microsoft Excel

Smoothing with Trend and Seasonality

Single exponential smoothing is intended for *trendless* data. If your data have a trend, you can try *Holt's method* with *two* smoothing constants (one for *trend,* one for *level*). If you have both trend and seasonality, you can try *Winters's method* with *three* smoothing constants (one for *trend,* one for *level,* one for *seasonality*). These advanced methods are similar to single smoothing in that they use simple formulas to update the forecasts, and you may use them without special caution. These topics are usually reserved for a class in forecasting, so they will not be explained here.

Mini Case 14.2

Exchange Rates

We have data for March 1 to March 30 and want to forecast 1 day ahead to March 31 by using exponential smoothing. We choose a smoothing constant value of α = .20 and set the initial forecast F_1 to the average of the first six actual values. Table 14.13 shows the

Table 14.13 Exchange Rate Canada/U.S. Dollar Canada

t	Date	Actual y_t	Forecast F_t	Error $e_t = y_t - F_t$
1	01-Mar	1.2425	1.23450	0.0080
2	02-Mar	1.2395	1.23610	0.0034
3	03-Mar	1.2463	1.23678	0.0095
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
21	29-Mar	1.2135	1.21406	− 0.0006
22	30-Mar	1.2164	1.21395	0.0024
23	31-Mar		**1.21444**	

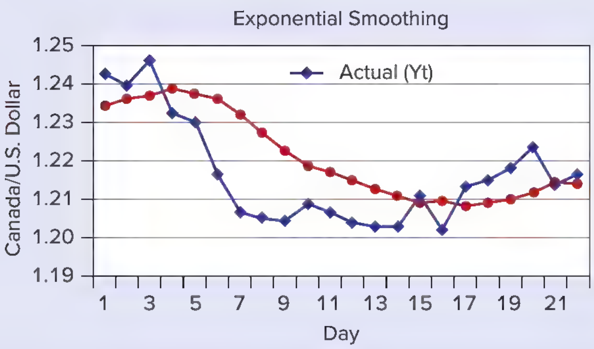
Source: Data from www.federalreserve.gov.

actual data (y_t) and forecasts (F_t) for each date. The March 31 forecast is $F_{23} = \alpha y_{22} + (1 - \alpha)F_{22} = (.20)(1.2164) + (.80)(1.21395) = 1.2144$.

The column of errors (e_t) shown in Table 14.13 can be used to calculate measures of fit (e.g., *MAPE, MAD*). Figure 14.22 shows that the forecasts adapt, but always with a lag. Exponential smoothing is really a kind of moving average, so its "forecasts" are mainly of short-term value.

Figure 14.22

Excel's Exponential Smoothing (α = .20)



Section Exercise

connect

14.7 (a) Make an Excel line graph of the following bond yield data (only the first and last three data values are shown). Describe the pattern. Is there a consistent trend? (b) Perform exponential smoothing with $\alpha = .20$. Use both methods *A* and *B* to initialize the forecast. Record the statistics of fit. (c) Do the smoothing again with $\alpha = .10$ and then with $\alpha = .30$, recording the statistics of fit. (d) Compare the statistics of fit for the three values of α. (e) Make a one-period forecast (i.e., $t = 53$) using each of the three α values. How did α affect your forecasts? **BondYield**

U.S. Treasury 10-Year Bond Yields at Week's End (n = 51 weeks)

Date	4-Jan	11-Jan	18-Jan	. . .	13-Dec	20-Dec	27-Dec
Rate	2.67	2.71	2.79	. . .	1.82	1.92	1.88

14.6 SEASONALITY

When and How to Deseasonalize

LO 14-7

Interpret seasonal factors and use them to make forecasts.

When the data periodicity is monthly or quarterly, we should calculate a seasonal index and use it to **deseasonalize** the data (annual data have no seasonality). This process is called **decomposition** of a time series. For a multiplicative model (the usual assumption), a seasonal index is a *ratio*. For example, if the seasonal index for July is 1.25, it means that July is 125 percent of the monthly average. If the seasonal index for January is 0.84, it means that January is 84 percent of the monthly average. If the seasonal index for October is 1.00, it means that October is an average month. The seasonal indexes must sum to 12 for monthly data or 4 for quarterly data. The following steps are used to deseasonalize data for time-series observations:

- Step 1 Calculate a centered moving average (*CMA*) for each month (quarter).
- Step 2 Divide each observed y_t value by the *CMA* to obtain seasonal ratios.
- Step 3 Average the seasonal ratios by month (quarter) to get raw seasonal indexes.
- Step 4 Adjust the raw seasonal indexes so they sum to 12 (monthly) or 4 (quarterly).
- Step 5 Divide each y_t by its seasonal index to get deseasonalized data.

In step 1, we lose 12 observations (monthly data) or 4 observations (quarterly data) because of the centering process. We will illustrate this technique for quarterly data.

Illustration of Calculations

Table 14.14 shows six years' data on quarterly revenue from sales of carpeting, tile, wood, and vinyl flooring by a floor-covering retailer. The data have an upward trend (see Figure 14.23), perhaps due to a boom in consumer spending on home improvement and new homes. There also appears to be seasonality, with lower sales in the third quarter (summer) and higher sales in the first quarter (winter). Excel has no seasonal decomposition feature, but you can perform your own calculations as shown in Table 14.15.

Table 14.14

Sales of Floor Covering Materials ($ thousands) FloorSales

Quarter	2012	2013	2014	2015	2016	2017
1	259	306	379	369	515	626
2	236	300	262	373	373	535
3	164	189	242	255	339	397
4	222	275	296	374	519	488

Table 14.15

Deseasonalized Sales (n = 24 quarters) FloorSales

Obs	Year	Quarter	Sales	CMA	Sales/CMA	Seasonal Index	Deseasonalized
1	2012	1	259			1.252	206.9
2		2	236			1.021	231.1
3		3	164	***226.125***	0.725	0.740	221.7
4		4	222	240.000	0.925	0.987	224.9
5	2013	1	306	251.125	1.219	1.252	244.4
6		2	300	260.875	1.150	1.021	293.8
7		3	189	276.625	0.683	0.740	255.5
8		4	275	281.000	0.979	0.987	278.6
9	2014	1	379	282.875	1.340	1.252	302.7
10		2	262	292.125	0.897	1.021	256.6
11		3	242	293.500	0.825	0.740	327.2
12		4	296	306.125	0.967	0.987	299.8
13	2015	1	369	321.625	1.147	1.252	294.7
14		2	373	333.000	1.120	1.021	365.3
15		3	255	361.000	0.706	0.740	344.7
16		4	374	379.250	0.986	0.987	378.8
17	2016	1	515	389.750	1.321	1.252	411.3
18		2	373	418.375	0.892	1.021	365.3
19		3	339	450.375	0.753	0.740	458.3
20		4	519	484.500	1.071	0.987	525.7
21	2017	1	626	512.000	1.223	1.252	500.0
22		2	535	515.375	1.038	1.021	524.0
23		3	397			0.740	536.7
24		4	488			0.987	494.3

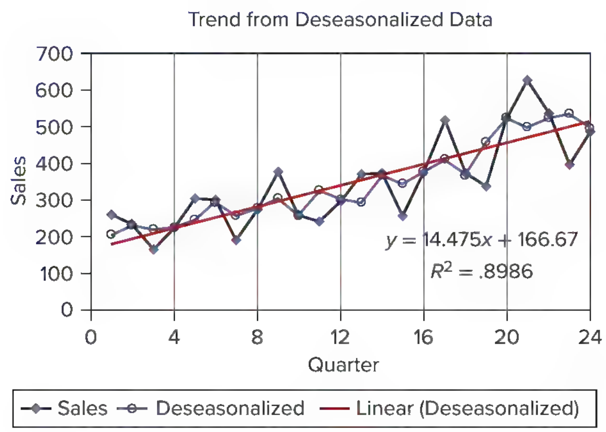


Figure 14.23

Deseasonalized Trend FloorSales

Table 14.16

Calculation of Seasonal Indexes FloorSales

Quarter	2012	2013	2014	2015	2016	2017	Mean	Adjusted
1		1.219	1.340	1.147	1.321	1.223	1.250	1.252
2		1.150	0.897	1.120	0.892	1.038	1.019	1.021
3	0.725	0.683	0.825	0.706	0.753		0.738	0.740
4	0.925	0.979	0.967	0.986	1.071		0.986	0.987
							3.993	4.000

Due to rounding, details may not yield the result shown.

Because the number of subperiods (quarters) is even ($m = 4$), each value of the *CMA* is the average of two averages. For example, the first *CMA* value 226.125 is the average of (259 + 236 + 164 + 222)/4 and (236 + 164 + 222 + 306)/4. Table 14.16 shows how the indexes are averaged. The *CMA* loses two quarters at the beginning and two quarters at the end, so each seasonal index is an average of only five quarters (instead of six). Each mean is then adjusted to force the sum to be 4.000, and these are the seasonal indexes. If we had monthly data, the indexes would be adjusted so that their sum would be 12.000. Calculations are ordinarily performed with software (e.g., MegaStat, Minitab, or R).

After the data have been deseasonalized, the trend (Figure 14.23) is fitted based on deseasonalized data. The sharper peaks and valleys in the original time series (Y) have been smoothed by removing the seasonality (S). Any remaining variation about the trend (T) is irregular (I) or "random noise." To make a forecast k periods ahead, multiply the *deseasonalized* trend estimate $\hat{y}_{t+k}$ by the seasonal index for period $t+k$.

Seasonal Forecasts Using Binary Predictors

LO 14-8

Use regression with seasonal binaries to make forecasts.

Another way to address seasonality is to estimate a regression model using **seasonal binaries** as predictors. For quarterly data, for example, the data set would look as shown in Table 14.17. When we have four binaries (i.e., four quarters), we must exclude one binary to prevent perfect multicollinearity (see Chapter 13, Section 13.5). Arbitrarily, we exclude the fourth quarter binary *Qtr4* (it will be a portion of the intercept when $Qtr1 = 0$ and $Qtr2 = 0$ and $Qtr3 = 0$).

Table 14.17

Sales Data with Seasonal Binaries FloorSales

Year	Quarter	Sales	Time	Qtr1	Qtr2	Qtr3
2012	1	259	1	1	0	0
	2	236	2	0	1	0
	3	164	3	0	0	1
	4	222	4	0	0	0
2013	1	306	5	1	0	0
	2	300	6	0	1	0
	3	189	7	0	0	1
	4	275	8	0	0	0
2014	1	379	9	1	0	0
	2	262	10	0	1	0
	3	242	11	0	0	1
	4	296	12	0	0	0
2015	1	369	13	1	0	0
	2	373	14	0	1	0
	3	255	15	0	0	1
	4	374	16	0	0	0
2016	1	515	17	1	0	0
	2	373	18	0	1	0
	3	339	19	0	0	1
	4	519	20	0	0	0
2017	1	626	21	1	0	0
	2	535	22	0	1	0
	3	397	23	0	0	1
	4	488	24	0	0	0

We assume a linear trend, and specify the regression model *Sales* = *f*(*Time, Qtr1, Qtr2, Qtr3*). The estimated regression is shown in Figure 14.24. This is an additive model of the form $Y = T + S + I$ (recall that we omit the cycle C in practice). The fitted equation (rounded) is

$$Sales = 161 + 14.4\ Time + 89.8\ Qtr1 + 12.9\ Qtr2 - 83.6\ Qtr3$$

Figure 14.24

Fitted Regression for Seasonal Binaries

	Coefficient	Standard Error	t Stat	p-Value
Intercept	161.208	24 334	6.625	0.0000
Time	14 366	1.244	11.549	0.0000
Qtr1	89.765	24.324	3.690	0.0016
Qrt2	12.899	24.164	0.534	0.5997
Qrt3	–83.634	24.068	–3.475	0.0025

$R^2 = .9001$ $SE = 41.63$ $F = 42.78$

Time is a significant predictor ($p = .0000$), indicating significant linear trend. Two of the binaries are significant: *Qtr1* ($p = .0016$) and *Qtr3* ($p = .0025$). The second-quarter binary *Qtr2* ($p = .5997$) is not significant. The model gives a good overall fit ($R^2 = .9001$). The main virtue of the seasonal regression model is its versatility. We can plug in future values of *Time* and the seasonal binaries to create forecasts as far ahead as we wish. For example:

$$\text{Period 25: } Sales = 161 + 14.4(25) + 89.8(1) + 12.9(0) - 83.6(0) = 610.8$$
$$\text{Period 26: } Sales = 161 + 14.4(26) + 89.8(0) + 12.9(1) - 83.6(0) = 548.3$$
$$\text{Period 27: } Sales = 161 + 14.4(27) + 89.8(0) + 12.9(0) - 83.6(1) = 466.2$$
$$\text{Period 28: } Sales = 161 + 14.4(28) + 89.8(0) + 12.9(0) - 83.6(0) = 564.2$$

Section Exercises

connect

14.8 (a) Plot the PepsiCo data. Is there a trend? (b) Do you see evidence of seasonality? (c) Use software of your choice (e.g., MegaStat, Minitab, or R) to deseasonalize the data and calculate quarterly seasonal indexes. (d) If there is seasonality, suggest possible reasons. (e*) Perform a regression using seasonal binaries. Interpret the results. **PepsiCo**

PepsiCo Revenues ($ billions), 2014–2019

Quarter	2014	2015	2016	2017	2018	2019
Qtr1	12.62	12.22	11.86	12.05	12.56	12.88
Qtr2	16.89	15.92	15.40	15.71	16.09	16.45
Qtr3	17.22	16.33	16.03	16.24	16.48	17.19
Qtr4	19.95	18.58	19.52	19.53	19.52	20.64

Source: Form 10-K reports for PepsiCo, Inc. and online earnings announcements. Data are for December 31 of each year.

14.9 (a) Plot the Corvette data. Is there a trend? (b) Do you see evidence of seasonality? (c) Use software of your choice (e.g., MegaStat, Minitab, or R) to deseasonalize the data and calculate monthly seasonal indexes. (d) If there is seasonality, suggest possible reasons. (e*) Perform a regression using seasonal binaries. Interpret the results. **Corvette**

U.S. Corvette Sales, 2004–2007 (number of cars sold)

Month	2004	2005	2006	2007
Jan	2,986	2,382	2,579	2,234
Feb	2,382	2,365	3,058	2,784
Mar	3,033	3,215	3,655	3,158
Apr	3,169	3,177	3,516	3,227
May	3,420	3,078	3,317	3,300
Jun	3,398	2,417	2,938	3,055
Jul	3,492	1,872	2,794	2,377
Aug	2,067	2,202	2,990	2,877
Sep	3,705	2,372	3,056	2,837
Oct	2,607	2,981	2,761	2,484
Nov	2,120	3,157	2,773	2,438
Dec	2,897	3,271	3,081	2,914
Total	35,276	32,489	36,518	33,685

Source: *Ward's Automotive Yearbook, 2005-2008.*

Mini Case 14.3

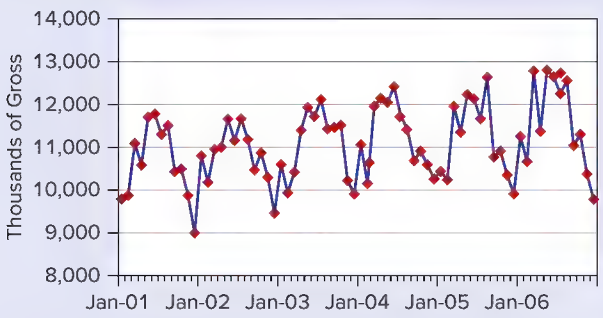
Using Seasonal Binaries Beer

Figure 14.25 shows monthly U.S. shipments of bottled beer for six years. A strong seasonal pattern is evident, presumably because people drink more beer in the warmer months. How can we describe the pattern statistically?

We create a regression data set with linear trend (Time = 1, 2, . . . , 72) and 11 seasonal binaries (Feb–Dec). The January binary is omitted to prevent perfect multicollinearity. The regression results, shown in Figure 14.26, indicate a good fit ($R^2 = .857$), significant upward trend ($p = 0.000$ for Time), and several seasonal binaries that differ significantly from zero (p-values near zero). Binary predictor coefficients indicate that shipments are above the January average during the spring and summer (Mar–Aug), below the January average in the winter (Nov–Feb), and near the January average in the fall (Sep–Oct). The fitted regression equation can be used to forecast any future months' shipments.

Figure 14.25

U.S. Bottled Beer Shipments, 2001–2006



Source: www.census.gov.

Figure 14.26

Minitab's Fitted Regression for Seasonal Binaries

The regression equation is
Beer = 10164 + 16.9 Time − 484 Feb + 768 Mar + 579 Apr
+ 1311 May + 1182 Jun + 975 Jul + 892 Aug − 99 Sep
+ 107 Oct − 644 Nov − 1089 Dec

Predictor	Coef	SE Coef	T	P
Constant	10163.7	161.4	62.97	0.000
Time	16.9	2.1	8.11	0.000
Feb	−483.9	209.2	−2.31	0.024
Mar	767.5	209.2	3.67	0.001
Apr	579.0	209.3	2.77	0.008
May	1310.7	209.3	6.26	0.000
Jun	1182.0	209.4	5.64	0.000
Jul	975.3	209.5	4.65	0.000
Aug	892.2	209.7	4.26	0.000
Sep	−99.3	209.8	−0.47	0.638
Oct	106.9	210.0	0.51	0.613
Nov	−644.1	210.2	−3.06	0.003
Dec	−1089.0	210.4	−5.18	0.000

S = 362.302 R-Sq = 85.7% R-Sq (adj) = 82.8%

14.7 INDEX NUMBERS

LO 14-9

Interpret index numbers.

A simple way to measure changes over time (and especially to compare two or more variables) is to convert time-series data into **index numbers**. The idea is to create an index that starts at 100 in a *base period,* so we can see *relative changes* in the data regardless of the original data units. Indexes are most often used for financial data (e.g., prices, wages, costs) but can be used with any numerical data (e.g., number of units sold, warranty claims, computer spam).

Relative Indexes

To convert a time series $y_1, y_2, \ldots, y_n$ into a *relative index* (sometimes called a *simple index*), we divide each data value y_t by the data value y_1 in a base period and multiply by 100. The relative index I_t for period t is

$$I_t = 100 \times \frac{y_t}{y_1} \tag{14.15}$$

The index in the base period is always $I_1 = 100$, so the index $I_1, I_2, \ldots, I_n$ makes it easy to see *relative changes* in the data, regardless of the original data units. For example, Table 14.18 shows 60 days of daily U.S. dollar exchange rates (on the left) and the corresponding index numbers (on the right) using November 1, 2019 = 100 as a base period. Because each index starts at the same point (100), we can easily see fluctuations and trends in Figure 14.27. We could fit a moving average, if we wanted to smooth the data. Speculators who engage in currency arbitrage would use even more-sophisticated tools to analyze movements in currency indexes.

Table 14.18

U.S. Foreign Exchange Rates, Currency

	Foreign Currency per Dollar			Index Numbers (Nov 1, 2019 = 100)		
Date	**U.K.**	**Mexico**	**Canada**	**U.K.**	**Mexico**	**Canada**
01-Nov-19	1.3145	19.0990	1.2950	100.0	100.0	100.0
04-Nov-19	1.3145	19.1660	1.2906	100.0	100.4	99.7
05-Nov-19	1.3176	19.2190	1.2870	100.2	100.6	99.4
⋮	⋮	⋮	⋮	⋮	⋮	⋮
28-Jan-20	1.3174	18.7800	1.2996	100.2	98.3	100.4
29-Jan-20	1.3201	18.6762	1.3012	100.4	97.8	100.5
30-Jan-20	1.3216	18.7990	1.3106	100.5	98.4	101.2

Source: Federal Reserve Bank of St. Louis (https://fred.stlouisfed.org).

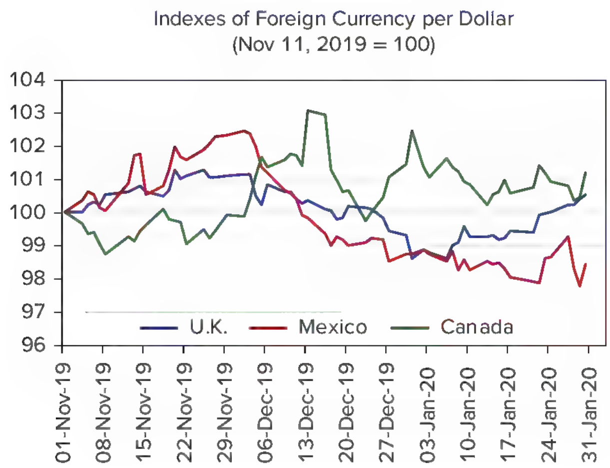


Figure 14.27

U.S. Foreign Exchange Rates, Currency

Weighted Indexes

A different calculation is required for a *weighted index* such as the Consumer Price Index for all urban consumers (CPI-U). The CPI-U is a measure of the relative prices paid by urban consumers for a market basket of goods and services, based on prices of hundreds of goods and services in eight major groups. The goal is to make the CPI-U representative of the prices paid for all goods and services purchased by all urban consumers. This requires assigning weights to each consumer good or service to reflect its importance relative to all the other goods and services in the market basket (e.g., housing gets a higher weight because it is a larger proportion of total spending). The basic formula for a simple weighted price index is

(14.16) $$I_t = 100 \times \frac{\sum_{i=1}^{m} p_{it} q_i}{\sum_{i=1}^{m} p_{i1} q_i} = 100 \times \frac{p_{1t}q_1 + p_{2t}q_2 + \cdots + p_{mt}q_m}{p_{11}q_1 + p_{21}q_2 + \cdots + p_{m1}q_m}$$

where

I_t = weighted index for period t ($t = 1, 2, \ldots, n$)

p_{it} = price of good i in period t ($i = 1, 2, \ldots, m$; $t = 1, 2, \ldots, n$)

q_i = weight assigned to good i ($i = 1, 2, \ldots m$)

The numerator is the cost of buying a given market basket of goods and services at today's prices (period t) relative to the cost of the same market basket in the base period (period 1). The weight q_i represents the relative quantity of the item in the consumer's budget. For example, suppose there is a price increase of 5 percent for food and beverages and a 10 percent increase for medical care costs, with no price changes for the other expenditure categories. This would result in an increase of 1.4 percent in the CPI, as shown in Table 14.19.

From Table 14.19, the price index rose from 100.0 to 101.4, or a 1.4 percent increase:

$$I_2 = 100 \times \frac{\sum_{i=1}^{m} p_{i2} q_i}{\sum_{i=1}^{m} p_{i1} q_i} = 100 \times \frac{101.4}{100.0} = 101.4$$

Formula (14.15) is called a *Laspeyres index*. It treats the base year quantity weights as constant. Weights are based on the *Survey of Consumer Expenditures*. In your economics classes, you may learn more sophisticated methods that take into account the fact that expenditure weights do change over time. One such method is the *Paasche index,* which uses a formula similar to the Laspeyres index, except that quantity weights are adjusted for each period.

Table 14.19

Illustrative Calculation of Price Index

Expenditure Category	Base Year ($t = 1$) Weight (q_i)		Relative Price (p_{i1})		Relative Spending ($p_{i1}q_i$)	Current Year ($t = 2$) Weight (q_i)		Relative Price (p_{i2})		Relative Spending ($p_{i2}q_i$)
Food and beverages	15.7	×	1.00	=	15.7	15.7	×	1.05	=	16.5
Housing	40.9	×	1.00	=	40.9	40.9	×	1.00	=	40.9
Apparel	4.4	×	1.00	=	4.4	4.4	×	1.00	=	4.4
Transportation	17.1	×	1.00	=	17.1	17.1	×	1.00	=	17.1
Medical care	5.8	×	1.00	=	5.8	5.8	×	1.10	=	6.4
Recreation	6.0	×	1.00	=	6.0	6.0	×	1.00	=	6.0
Education/communication	5.8	×	1.00	=	5.8	5.8	×	1.00	=	5.8
Other goods and services	4.3	×	1.00	=	4.3	4.3	×	1.00	=	4.3
Sum	100.0		$\sum_{i=1}^{m} p_{i1} q_i$	=	100.0	100.0		$\sum_{i=1}^{m} p_{i2} q_i$	=	101.4

Importance of Index Numbers

The CPI affects nearly all Americans because it is used to adjust things like retirement benefits, food stamps, school lunch benefits, alimony, and tax brackets. The CPI-U could be compared with an index of salary growth for workers, or to measure current-dollar salaries in "real dollars." The Bureau of Labor Statistics (www.bls.gov) publishes CPI historical statistics for 31 categories. The most widely used CPI-U uses 1982–84 as a reference. That is, the Bureau of Labor Statistics sets the CPI-U (the average price level) for the years 1982, 1983, and 1984 equal to 100, and then measures changes in relation to that figure. As of December 2019 for example, the CPI-U was 257.0 meaning that, on average, prices had more than doubled over the previous 36 years (about a 2.7 percent annual increase, applying the geometric mean formula 4.5 with $n = 36$). The CPI is based on the buying habits of the "average" consumer, so it may not be a perfect reflection of anyone's individual price experience.

Other familiar price indexes, such as the Dow Jones Industrial Average (DJIA), have their own unique methodologies. Originally a simple arithmetic mean of stock prices, the DJIA now is the sum of the 30 stock prices divided by a "divisor" to compensate for stock splits and other changes over time. The divisor is revised periodically. Because high-priced stocks comprise a larger proportion of the sum, the DJIA is more strongly affected by changes in high-priced stocks. A little web research can tell you a lot about how stock price indexes are calculated, their strengths and weaknesses, and some alternative indexes that finance experts have invented.

14.8 FORECASTING: FINAL THOUGHTS

Role of Forecasting

In many ways, forecasting resembles planning. *Forecasting* is an analytical way to describe a "what-if" future that might confront the organization. *Planning* is the organization's attempt to determine actions it will take under each foreseeable contingency. Forecasts help decision makers become aware of trends or patterns that require a response. Actions taken by the decision makers may actually head off the contingency envisioned in the forecast. Thus, forecasts tend to be self-defeating because they trigger homeostatic organizational responses. So-called "black swan" events (https://en.wikipedia.org/wiki/Black_swan_theory) such as the COVID-19 crisis pose a vast challenge in making projections of many time series variables.

Behavioral Aspects of Forecasting

Forecasts can facilitate organizational communication. The forecast (or even just a nicely prepared time-series chart) lets everyone examine the same facts concurrently and perhaps argue with the data or the assumptions that underlie the forecast or its relevance to the organization. A quantitative forecast helps *make assumptions explicit.* Those who prepare the forecast must explain and defend their assumptions, while others must challenge them. In the process, everyone gains understanding of the data, the underlying realities, and the imperfections in the data. Forecasts *focus the dialogue* and can make it more productive.

Of course, this assumes a certain maturity among the individuals around the table. Strong leaders (or possibly meeting facilitators) can play a role in guiding the discourse to produce a positive result. The danger is that people may try to find scapegoats (yes, they do tend to blame the forecaster), deny facts, or avoid responsibility for tough decisions. But one premise of this book is that statistics, when done well, can strengthen any dialogue and lead to better decisions.

Forecasts Are Always Wrong

We discussed several measures to use to determine if a forecast model fits the time series. Successful forecasters understand that a forecast is never precise. There is always some error, but we can *use* the error measures to track forecast error. Many companies use several different forecasting models and rely on the model that has had the least error over some time period. We have described simple models in this chapter. You may take a class specifically focusing on forecasting in which you will learn about other time-series models including AR

(autoregressive) and ARIMA (autoregressive integrated moving average) models. Such models take advantage of the dependency that might exist between values in the time series. To ensure good forecast outcomes

- Maintain up-to-date databases of *relevant* data.
- Allow sufficient lead time to analyze the data.
- State several alternative forecasts or scenarios.
- Track forecast errors over time.
- State your assumptions and qualifications and consider your time horizon.
- Don't underestimate the power of a good graph.

Mini Case 14.4

How Does Noodles & Company Ensure Its Ingredients Are as Fresh as Possible?

Using only fresh ingredients is key for great food and success for restaurants like Noodles & Company. To be sure that the restaurants are serving only the freshest ingredients, while also reducing food waste, Noodles & Company turned to statistical forecasting for ordering ingredients and daily food preparation. The challenge was to create a forecast that is sophisticated enough to be accurate yet simple enough for new restaurant employees to understand.

Noodles & Company uses a food management software system to forecast the demand for its menu items based on the moving average of the previous four weeks' sales. This simple forecasting technique has been very accurate. The automated process also uses the forecast of each item to estimate how many ingredients to order as well as how much to prepare each day. For example, the system might forecast that during next Wednesday's lunch, the location in Longmont, Colorado, will sell 55 Pesto Cavatappis. After forecasting the demand for each menu item, the system then specifies exactly how much of each ingredient to prepare for that lunch period.

For the restaurant teams, the old manual process of estimating and guessing how much of each ingredient to prepare is now replaced with an automated prep sheet. Noodles & Company has reduced food waste because restaurants are less likely to overorder ingredients and overprepare menu items. The restaurant teams are more efficient and customers are served meals made with the freshest ingredients possible.

Chapter Summary

A **time series** is assumed to have four components. For most business data, **trend** is the general pattern of change over all years observed, while **cycle** is a repetitive pattern of change around the trend over several years and **seasonality** is a repetitive pattern within a year. The **irregular** component is a random disturbance that follows no pattern. The **additive model** is adequate in the short run because the four components' magnitude does not change much, but for observations over longer periods of time, the **multiplicative model** is preferred. Common trend models include **linear** (constant slope and no turning point), **quadratic** (one turning point), and **exponential** (constant percent growth or decline). Higher polynomial models are untrustworthy and liable to give strange forecasts, though any trend model is less reliable the farther out it is projected. In forecasting, forecasters use fit measures besides R^2, such as mean absolute percent error (***MAPE***), mean absolute deviation (***MAD***), and mean squared deviation (***MSD***). For trendless or erratic data, we use a **moving average** over m periods or **exponential smoothing.** Forecasts adapt rapidly to changing data when the **smoothing constant** α is large (near 1) and conversely for a small α (near 0). For monthly or quarterly data, a **seasonal adjustment** is required before extracting the trend. Alternatively, regression with **seasonal binaries** can be used to capture seasonality and make forecasts. **Index numbers** are used to show changes relative to a base period.

Key Terms

additive model
centered moving average (*CMA*)
coefficient of determination
cycle
decomposition
deseasonalize
exponential smoothing
exponential trend
index numbers
irregular
linear trend
mean absolute deviation (*MAD*)
mean absolute percent error (*MAPE*)
mean squared deviation (*MSD*)
moving average
multiplicative model
periodicity
polynomial model
quadratic trend
seasonal
seasonal binaries
smoothing constant
standard error (SE)
time-series variable
trailing moving average (*TMA*)
trend

Commonly Used Formulas

Additive time-series model: $Y = T + C + S + I$

Multiplicative time-series model: $Y = T \times C \times S \times I$

Linear trend model: $y_t = a + bt$

Exponential trend model: $y_t = ae^{bt}$

Quadratic trend model: $y_t = a + bt + ct^2$

Coefficient of determination: $$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

Mean absolute percent error: $$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t}$$

Mean absolute deviation: $$MAD = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

Mean squared deviation: $$MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$$

Standard error: $$SE = \sqrt{\sum_{t=1}^{n}\frac{(y_t - \hat{y}_t)^2}{n-2}}$$

Forecast updating equation for exponential smoothing: $F_{t+1} = \alpha y_t + (1 - \alpha)F_t$

Chapter Review

1. Explain the difference between (a) stocks and flows; (b) cross-sectional and time-series data; (c) additive and multiplicative models.
2. (a) What is periodicity? (b) Give original examples of data with different periodicity.
3. (a) What are the distinguishing features of each component of a time series (trend, cycle, seasonal, irregular)? (b) Why is cycle usually ignored in time-series modeling?
4. Name four criteria for assessing a trend forecast.
5. Name two advantages and two disadvantages of each of the common trend models (linear, exponential, quadratic).
6. When would the exponential trend model be preferred to a linear trend model?
7. Explain how to obtain the compound percent growth rate from a fitted exponential model.
8. (a) When might a quadratic model be useful? (b) What precautions must be taken when forecasting with a quadratic model? (c) Why are higher-order polynomial models dangerous?
9. Name five measures of fit for a trend, and state their advantages and disadvantages.
10. (a) When do we use a moving average? (b) Name two types of moving averages. (c) When is a centered moving average harder to calculate?
11. (a) When is exponential smoothing most useful? (b) Interpret the smoothing constant α. What is its range? (c) What does a small α say about the degree of smoothing? A large α?
12. (a) Explain two ways to initialize the forecasts in an exponential smoothing process. (b) Name an advantage and a disadvantage of each method.
13. (a) Why is seasonality irrelevant for annual data? (b) List the steps in deseasonalizing a monthly time series. (c) What is the sum of a monthly seasonal index? A quarterly index?
14. (a) How can forecasting improve communication within an organization? (b) List five tips for ensuring effective forecasting outcomes.

15. (a) Explain how seasonal binaries can be used to model seasonal data. (b) What is the advantage of using seasonal binaries?

16. Explain the equivalency between the two forms of an exponential trend model.

17. What is the purpose of index numbers?

Chapter Exercises

Instructions: For each exercise, make an attractive, well-labeled time-series line chart. Adjust the *Y*-axis scale if necessary to show more detail (because Excel usually starts the scale at zero). If a fitted trend is called for, display the equation and R^2 statistic (or *MAPE, MAD,* and *MSD* in Minitab). Include printed copies of all relevant graphs with your answers to each exercise. Exercises marked with * are based on harder material.

14.10 (a) Make a line chart for JetBlue's revenue. (b) Describe the trend (if any) and discuss possible causes. (c) Fit both a linear and an exponential trend to the data. (d) Which model is preferred? Why? **JetBlue**

JetBlue Airlines Revenue, 2012–2019 (billions)

Year	Revenue
2012	4.98
2013	5.44
2014	5.82
2015	6.42
2016	6.63
2017	7.01
2018	7.66
2019	8.09

Source: JetBlue's published annual Form 10-K reports. Data are for December 31 of each year.

14.11 (a) Plot both Swiss watch time series on the same graph. (b) Describe the trend (if any) and discuss possible causes. (c) Fit an exponential trend to each time series. (d) Interpret each fitted trend carefully. What conclusion do you draw? **Swiss**

Swiss Watch Exports (thousands of units), 2014–2019

Year	Mechanical	Electronic
2014	8,131	20,455
2015	7,812	20,325
2016	6,963	18,433
2017	7,238	17,068
2018	7,525	16,215
2019	7,236	13,398

Source: Fédération de L'Industrie Horlogère Suisse, Swiss Watch Exports, https://www.fhs.swiss/eng/statistics.html

14.12 (a) Make a line graph of the U.S. civilian labor force data. (b) Describe the trend (if any) and discuss possible causes. (c) Fit three trend models: linear, exponential, and quadratic. Which model would offer the most believable forecasts? Explain. (d) Choose one of the fitted trend models and make forecasts for years 2020–2022. Justify your choice. **LaborForce**

U.S. Civilian Labor Force (thousands)

Year	Labor Force
2010	153,156
2011	153,373
2012	154,904
2013	154,408
2014	155,521
2015	157,245
2016	158,968
2017	159,880
2018	162,510
2019	164,007

Source: www.bls.gov. Data are in thousands for December 31 of each year, seasonally adjusted.

14.13 (a) Plot the voter participation rate. (b) Describe the trend (if any) and discuss possible causes. (c) Fit three trend models: linear, exponential, and quadratic. (d) Which trend model is preferred? Why? (e) Make a forecast for 2020, using a trend model of your choice (f*) *Optional challenge:* Check your forecast. How accurate was it? Discuss. *Note:* Time is in four-year increments, so use $t = 16$ for the 2020 forecast. **Voters**

U.S. Presidential Election Voter Participation, 1960–2016

Year	Voting Age Population	Voted for President	% Voting Pres
1960	109,672	68,838	62.8
1964	114,090	70,645	61.9
1968	120,285	73,212	60.9
1972	140,777	77,719	55.2
1976	152,308	81,556	53.5
1980	163,945	86,515	52.8
1984	173,995	92,653	53.3
1988	181,956	91,595	50.3
1992	189,524	104,425	55.1
1996	196,928	96,278	49.0
2000	207,884	105,397	50.7
2004	220,377	122,349	55.5
2008	229,945	131,407	57.1
2012	235,248	129,235	54.9
2016	251,107	138,885	54.4

Sources: *Statistical Abstract of the United States, 2011,* www.census.gov. Population and voters are in thousands.

14.14 For each of the following fitted trends, make a prediction for period $t = 17$:

a. $y_t = 2286e^{.076t}$
b. $y_t = 1149 + 12.78t$
c. $y_t = 501 + 18.2t - 7.1t^2$

14.15 (a) Make a line graph of total consumer credit outstanding. (b) Describe the trend (if any) and discuss possible causes. (c) Fit linear, exponential, and quadratic trends. (d) Plot both revolving and nonrevolving credit on the same graph. Do the trends differ? Explain. **Consumer**

14.16 (a) Plot the data on U.S. general aviation shipments. (b) Describe the pattern and discuss possible causes. *Hint:* What economic factors affect major capital investments? (c) Would a fitted trend be helpful? Explain. (d) Fit a moving average (e.g., period 2) to the data. Is it useful? (e) Would trend forecasts for 2020 and beyond be appropriate? Explain. **Airplanes**

14.17 For each of the following fitted trends, make a prediction for period $t = 12$:

a. $y_t = 372e^{-.041t}$
b. $y_t = 719 + 10t$
c. $y_t = 1299 - 51t + 7t^2$

14.18 (a) Plot *either* receipts and outlays *or* federal debt and GDP (plot both time series on the same graph). (b) Describe the two trends and discuss possible causes. (c) Fit an exponential trend to each. (d) Compare the growth rates. Explain the implications. (e) Plot the ratio of debt to GDP. What does it show? **FedBudget**

U.S. Consumer Credit Outstanding, 2000–2019 ($ billions) **Consumer**

Year	Total	Revolving	Nonrevolving
2000	1,722	683	1,039
2001	1,868	715	1,153
2002	1,972	751	1,221
.	.	.	.
.	.	.	.
.	.	.	.
2017	3,828	1,022	2,806
2018	4,010	1,054	2,956
2019	4,176	1,086	3,090

Source: www.federalreserve.gov. Data are for December 31 of each year. Units are billions of dollars, seasonally adjusted. Total is short and intermediate term credit to individuals, the sum of revolving credit (mostly credit card and home equity loans) and nonrevolving credit (for a specific purchase such as a car, mobile home, education, boats, trailers, or vacations).

U.S. Manufactured General Aviation Shipments, 2002–2019 **Airplanes**

Year	Planes	Year	Planes	Year	Planes	Year	Planes
2002	2,207	2007	3,279	2012	1,516	2017	1,595
2003	2,137	2008	3,079	2013	1,615	2018	1,746
2004	2,355	2009	1,585	2014	1,631	2019	1,771
2005	2,857	2010	1,334	2015	1,592		
2006	3,147	2011	1,323	2016	1,531		

Source: U.S. Manufactured General Aviation Shipments, *Statistical Databook,* General Aviation Manufacturers Association.

U.S. Federal Finances, 2001–2019 ($ billions current) **FedBudget**

Year	Receipts	Outlays	Fed Debt	GDP	Debt/GDP
2001	1,991	1,863	5,770	10,565	0.546
2002	1,853	2,011	6,198	10,877	0.570
2003	1,782	2,160	6,760	11,332	0.597
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
2017	3,316	3,982	20,206	19,272	1.048
2018	3,330	4,109	21,462	20,236	1.061
2019	3,462	4,447	22,668	21,220	1.068

Source: *Economic Report of the President, 2019.*

14.19 (a) Plot both men's and women's winning times (in minutes) on the same graph. (b) Fit a linear trend model to each series. Ask Excel for forecasts 20 years ahead. From the fitted trends, will the times eventually converge? Explain. (c) Make a copy of your graph, click each fitted trend, and change it to a moving average trend type. (d) Is a moving average a reasonable approach to modeling these data sets? *Note:* The data file **Boston** has the data converted to decimal minutes. Only the first three and last three lines are shown here.

14.20 (a) Plot the data on leisure and hospitality employment. (b) Describe the trend (if any) and discuss possible causes. (c) Fit the linear, exponential, and quadratic trends. Would any of these trend models give credible forecasts for 2020 and beyond? Explain. **Leisure**

Leisure and Hospitality Employment, 2006–2019 (thousands)

Year	Employees	Year	Employees
2006	13,292	2013	14,454
2007	13,550	2014	14,892
2008	13,256	2015	15,407
2009	12,944	2016	15,845
2010	13,158	2017	16,195
2011	13,538	2018	16,555
2012	13,978	2019	16,942

Source: http://data.bls.gov.

14.21 (a) Plot the data on law enforcement officers killed. (b) Describe the trend (if any) and discuss possible causes. (c) Would a fitted trend be helpful? Explain. **LawOfficers**

U.S. Law Enforcement Officers Killed, 2006–2016

Year	Killed	Year	Killed
2006	114	2012	97
2007	141	2013	76
2008	109	2014	96
2009	96	2015	86
2010	128	2016	135
2011	125	2017	128

Source: https://ucr.fbi.gov.

14.22 (a) Plot the data on lightning deaths. (b) Describe the trend (if any) and discuss possible causes. (c) Fit an exponential trend to the data. Interpret the fitted equation. (d) Make a forecast for 2020, using a trend model of your choice (or a judgment forecast). Explain the basis for your forecast. *Note:* Time is in five-year increments, so use $t = 17$ for your 2020 forecast. **Lightning**

U.S. Lightning Deaths, 1940–2010

Year	Deaths	Year	Deaths
1940	340	1980	74
1945	268	1985	74
1950	219	1990	74
1955	181	1995	85
1960	129	2000	51
1965	149	2005	38
1970	122	2010	29
1975	91	2015	27

Sources: *Statistical Abstract of the United States, 2011*, p. 234, and www.nws.noaa.gov.

Boston Marathon Champions, 1980–2019 **Boston**

Men			Women		
Year	Name of Winner	Time	Year	Name of Winner	Time
1980	Bill Rodgers	2:12:11	1980	Jacqueline Gareau	2:34:28
1981	Toshihiko Seko	2:09:26	1981	Allison Roe	2:26:46
1982	Alberto Salazar	2:08:52	1982	Charlotte Teske	2:29:33
. . .	. . .	. . .	. . .	. . .	. . .
2017	Geoffrey Kirui	2:09:37	2017	Edna Kiplagat	2:21:52
2018	Yuki Kawauchi	2:15:58	2018	Desi Linden	2:39:54
2019	Lawrence Cherono	2:07:57	2019	Worknesh Degefa	2:23:31

Source: www.wikipedia.org.

14.23 (a) Plot the data on skier/snowboard visits. (b) Would a fitted trend be helpful? Explain. **SnowBoards**

14.24 (a) Plot both men's and women's 100-meter dash winning times on the same graph. (b) Fit a linear trend model to each series (men, women). (c) Use Excel's option to forecast each trend graphically to 2040 (i.e., up to period $t = 27$ periods because observations are in four-year increments). From these projections, does it appear that the times will eventually converge? *Optional challenge:* (d*) Set the fitted trends equal, solve for x (the time period when the trends will cross), and convert x to a year. Is the result plausible? Explain. *Note:* Only the first three and last three years are displayed here. **Olympic**

14.25 (a) Plot U.S. petroleum imports on a graph. (b) Describe the trend (if any) and discuss possible causes. (c) Fit linear, exponential, and quadratic trends. (c) Do any of these trends seem appropriate to make forecasts? Explain. (d) Make a projection for 2020 using any method (including judgment or a moving average). Do you believe it? Explain. *Note:* Time increments are five years, so use $t = 13$ for the 2020 forecast. **Petroleum**

U.S. Annual Petroleum Imports, 1960–2015 (billions of barrels)

Year	Imports	Year	Imports
1960	372	1990	2,151
1965	452	1995	2,639
1970	483	2000	3,320
1975	1,498	2005	3,696
1980	1,926	2010	3,363
1985	1,168	2015	2,687

Source: www.eia.doe.gov.

14.26 (a) Make a line chart for an m-period moving average to the exchange rate data shown below (only the first 3 and last 3 days are shown). with $m = 2$, 3, 4, and 5 periods. For each method, state the last MA value. (b) Which value of m do you prefer? Why? (c) Is a moving average appropriate for this kind of data? **Sterling**

U.S. Skier/Snowboarder Visits, 2000–2019 (millions) **SnowBoards**

Season	Visits	Season	Visits	Season	Visits
2000–01	57.3	2007–08	60.5	2014–15	53.6
2001–02	54.4	2008–09	57.4	2015–16	52.8
2002–03	57.6	2009–10	59.8	2016–17	54.8
2003–04	57.1	2010–11	60.5	2017–18	53.3
2004–05	56.9	2011–12	51.0	2018–19	59.3
2005–06	58.9	2012–13	56.9		
2006–07	55.1	2013–14	56.5		

Source: www.nsaa.org/nsaa/press/industryStats.asp.

Summer Olympics 100-Meter Dash Winning Times, 1928–2016 **Olympic**

Year	Men's 100-Meter Winner	Seconds	Women's 100-Meter Winner	Seconds
1928	Percy Williams, Canada	10.80	Elizabeth Robinson, United States	12.20
1932	Eddie Tolan, United States	10.30	Stella Walsh, Poland	11.90
1936	Jesse Owens, United States	10.30	Helen Stephens, United States	11.50
...	...	....	...	...
2008	Usain Bolt, Jamaica	9.69	Shelly-Ann Fraser, Jamaica	10.78
2012	Usain Bolt, Jamaica	9.63	Shelly-Ann Fraser, Jamaica	10.75
2016	Usain Bolt, Jamaica	9.81	Elaine Thompson, Jamaica	10.71

Source: www.wikipedia.org.

Daily Spot Exchange Rate, U.S. Dollars per Pound Sterling ($n = 60$ days) **Sterling**

Day	1	2	3	. . .	58	59	60
Date	11/1/19	11/4/19	11/5/19	. . .	1/28/20	1/29/20	1/30/20
Rate	1.2950	1.2906	1.2870	. . .	1.2996	1.3012	1.3106

Source: Federal Reserve Board of Governors.

14.27 Refer to exercise 14.26. (a) Plot the dollar/pound exchange rate data. Copy and paste the chart so that you have four copies (one for each α). (b) Perform simple exponential smoothing (using Excel's Data Analysis or other software such as Minitab) using $\alpha = .05, .10, .20,$ and $.50$. (c) Which value of α do you prefer? Why? (d) Is exponential smoothing appropriate for this kind of data? **Sterling**

14.28 (a) Plot the data on natural gas bills. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or Minitab or R to calculate estimated seasonal indexes. (d) Which months are the highest? The lowest? Can you explain this pattern? (e) Is there a trend in the deseasonalized data? *Optional challenge:* (f*) Perform a regression using 11 seasonal binaries. Interpret the results. **GasBills**

14.29 (a) Plot the data on air travel delays. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or Minitab or R to calculate estimated seasonal indexes. (d) Which months have the most delays? The fewest? Can you suggest reasons? **Delays**

14.30 (a) Plot the data on airplane shipments. (b) Can you see seasonal patterns? Explain. (c) Use MegaStat or Minitab or R to calculate estimated seasonal indexes. (d) In which quarters are shipments highest? Lowest? Can you suggest reasons? **AirplanesQtr**

14.31 Ten years ago, on the last trading day of year 1, Felicia invested \$1,000 in each of three stock funds. On the last trading day of year 10, their values were:

Fund A: \$2,509 Fund B: \$2,096 Fund C: \$3,034.

Use formula 14.4 to estimate the implied rate of return for each fund.

14.32 (a) Use MegaStat or Minitab or R to deseasonalize the quarterly data on Coca-Cola's revenues and calculate seasonal indexes. (b) Interpret the seasonal indexes. If there is seasonality, suggest possible reasons. (c*) Perform a regression using seasonal binaries. Interpret the results. **CocaCola**

Natural Gas Bills for a California Residence, 2017–2020 **GasBills**

Month	2017	2018	2019	2020
Jan	78.98	118.86	101.44	155.37
Feb	84.44	111.31	122.20	148.77
Mar	65.54	75.62	99.49	115.12
Apr	62.60	77.47	55.85	85.89
May	29.24	29.23	44.94	46.84
Jun	18.10	17.10	19.57	24.93
Jul	91.57	16.59	15.98	20.84
Aug	6.48	27.64	14.97	26.94
Sep	19.35	28.86	18.03	34.17
Oct	29.02	48.21	56.98	88.58
Nov	94.09	67.15	115.27	100.63
Dec	101.65	125.18	130.95	174.63

Source: Homeowner's records.

U.S. Airspace Total System Delays, 2013–2017 **Delays**

Month	2013	2014	2015	2016	2017
Jan	16,240	15,385	18,571	18,035	29,548
Feb	17,031	19,755	18,553	20,989	25,607
Mar	21,697	20,227	22,326	28,237	38,291
Apr	37,117	25,912	24,416	22,683	41,977
May	35,740	35,218	31,125	28,455	49,208
Jun	46,693	43,059	41,560	39,238	52,981
Jul	46,715	37,967	38,308	43,881	49,913
Aug	31,101	34,499	32,711	41,335	47,951
Sep	21,844	28,302	25,455	27,085	32,091
Oct	21,066	31,940	21,893	26,619	31,248
Nov	16,316	20,647	21,376	23,498	20,732
Dec	21,809	28,206	29,087	25,411	25,381

Source: www.faa.gov.

14.33 (a) Use MegaStat or Minitab or R to deseasonalize the monthly data on student pilots and calculate seasonal indexes. (b) Interpret the seasonal indexes. If there is seasonality, suggest possible reasons. (c*) Perform a regression using seasonal binaries. Interpret the results. **StudentPilots**

14.34 The following seasonal regression was fitted with quarterly seasonal binaries beginning in the first quarter (*Qtr4* is omitted to avoid multicollinearity). Make a prediction for y_t in period (a) $t = 21$; (b) $t = 8$; (c) $t = 15$.

$$y_t = 213 + 11t - 9\,Qtr1 + 12\,Qtr2 - 15\,Qtr3.$$

14.35 The following seasonal regression was fitted with quarterly seasonal binaries beginning in the first quarter (*Qtr4* is omitted to avoid multicollinearity). Make a prediction for y_t in period (a) $t = 14$; (b) $t = 17$; (c) $t = 20$.

$$y_t = 491 + 19t + 29\,Qtr1 - 18\,Qtr2 + 12\,Qtr3.$$

14.36 You want to invest $1,000. Which growth curve would yield the largest principal 5 years from now? 10 years? 20 years? Explain. *Hint:* Show all the forecasts.

a. $y_t = 1000e^{0.021t}$
b. $y_t = 1000 + 25t$
c. $y_t = 1000 + 28t - 0.4t^2$

U.S. Manufactured General Aviation Shipments, 2012–2019 **AirplanesQtr**

Year	Qtr 1	Qtr 2	Qtr 3	Qtr 4	Total
2012	305	358	339	514	1,516
2013	329	413	353	520	1,615
2014	345	380	379	527	1,631
2015	296	374	378	544	1,592
2016	280	357	385	509	1,531
2017	311	379	377	532	1,599
2018	315	428	416	587	1,746
2019	366	401	423	581	1,771

Note: Quarterly shipments may not add to annual total because some manufacturers report only annual totals.
Source: U.S. Manufactured General Aviation Shipments, *Statistical Databook,* General Aviation Manufacturers Association, used with permission.
Source: "U.S. Manufactured General Aviation Shipments," *Statistical Databook,* General Aviation Manufacturers Association.
Note that quarterly shipments may not add to annual total because some manufacturers report only annual totals.

Coca-Cola Revenues ($ millions), 2014–2019 **CocaCola**

Quarter	2014	2015	2016	2017	2018	2019
Qtr1	10.58	10.71	10.28	9.12	7.63	8.02
Qtr2	12.57	12.16	11.54	9.70	9.42	10.00
Qtr3	11.98	11.43	10.63	9.08	8.78	9.51
Qtr4	10.87	10.00	9.41	7.51	5.36	9.07

Sources: 10-K reports of The Coca-Cola Company.

Student Pilot Certificates Issued by Month, 2013–2018 **StudentPilots**

Month	2013	2014	2015	2016	2017	2018
Jan	4,480	3,882	3,805	3,714	2,173	3,202
Feb	3,921	3,154	3,327	3,700	2,180	3,462
Mar	4,662	3,451	3,833	5,287	3,250	4,110
Apr	3,693	3,881	3,918	1,753	2,495	3,441
May	4,029	4,159	3,882	2,948	2,828	3,958
Jun	4,336	4,614	4,856	3,001	3,128	3,611
Jul	4,789	4,833	4,659	3,096	3,141	4,460
Aug	5,492	5,104	4,867	3,670	4,536	3,998
Sep	4,025	4,195	4,188	3,921	2,588	4,242
Oct	3,926	3,963	3,863	2,815	5,534	4,635
Nov	3,293	3,133	3,061	1,302	3,945	3,140
Dec	2,920	3,038	3,122	938	2,603	3,095

Source: www.faa.gov/data_statistics/aviation_data_statistics.

Related Reading

Hanke, John E., and Dean W. Wichern. *Business Forecasting*. 9th ed. Pearson, 2014.

Ord, Keith, and Robert Fildes. *Principles of Forecasting*. Cengage, 2013.

Wilson, J. Holton, and Barry Keating. *Forecasting and Predictive Analytics with Forecast X™*, 7th ed. McGraw-Hill, 2019.

CHAPTER 14 More Learning Resources

You can access these *LearningStats* demonstrations through McGraw-Hill's Connect® to help you understand time-series analysis.

Topic	*LearningStats Demonstrations*
Trends and forecasting	Trend Forecasting Measures of Fit Exponential Trend Formula
Simulations	Time-Series Components Trend Simulator Seasonal Time-Series Generator
Exponential smoothing	Advanced Forecasting Methods Single Exponential Smoothing Brown's Double Smoothing Holt-Winters Seasonal Smoothing Exponential Smoothing Weights
ARIMA Models	ARIMA Terminology ARIMA Patterns ARIMA Calculations Seasonal ARIMA

Key: = PowerPoint = Excel = Adobe PDF

Software Supplement

Minitab Trend Analysis

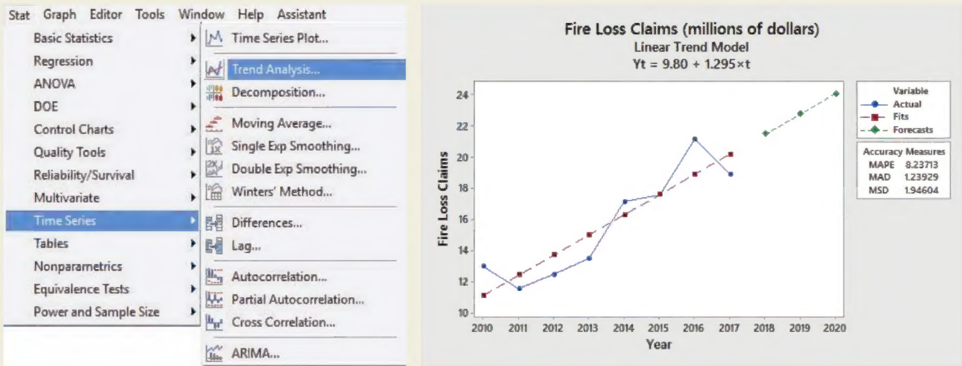
Figure 14.28 shows Minitab's trend menu and time plot of fire loss claims. Minitab displays fit statistics *MAPE, MAD,* and *MSD* instead of R^2. An attractive feature is Minitab's separation between actual and forecasts, and using different color for actual (blue) and forecasts (green).

Figure 14.28

Minitab's Trend Analysis

FireLosses

Source: Minitab



Minitab Exponential Smoothing

Figure 14.29 shows Minitab's single exponential smoothing and four weeks' forecasts for the deck sealer data (Excel's exponential smoothing does not make forecasts). After week 18, the exponential smoothing method cannot be updated with actual data, so the forecasts are constant. The wide 95 percent confidence intervals (red triangles) reflect the erratic past sales pattern.

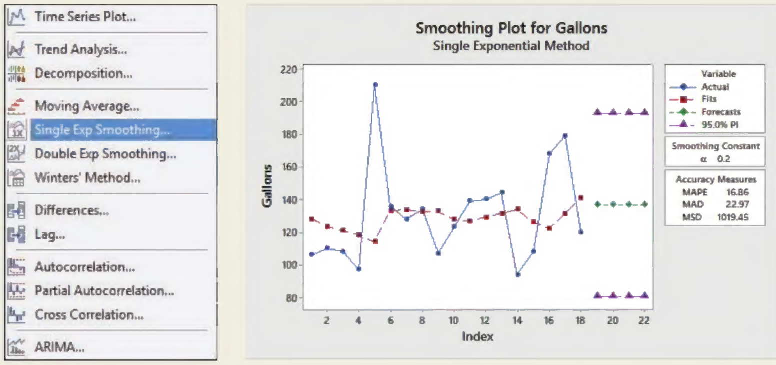


Figure 14.29

Minitab's Exponential Smoothing **DeckSealer**

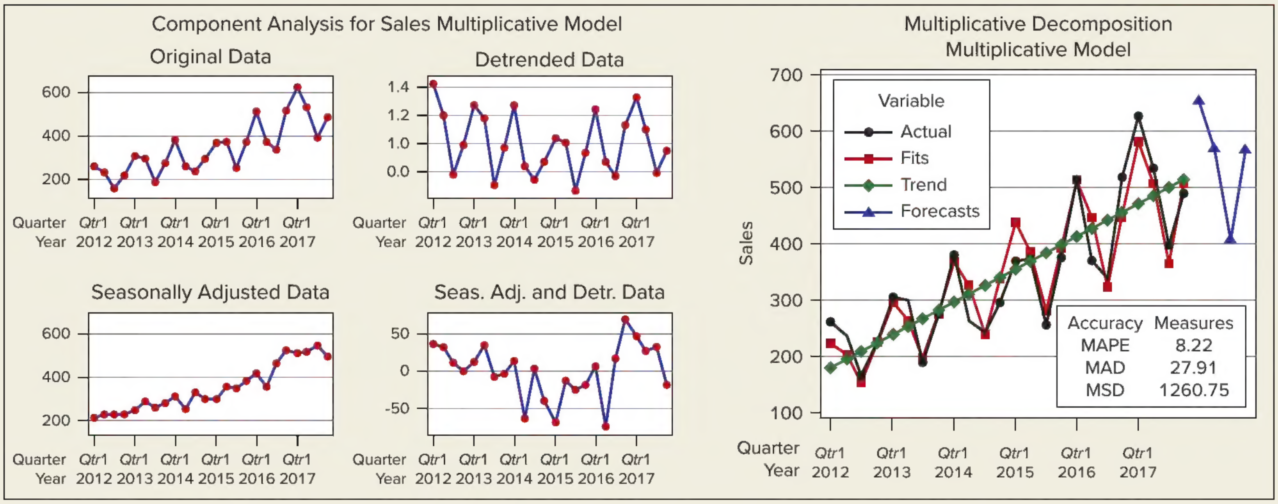
Source: Minitab

Using Minitab to Deseasonalize

Minitab performs its deseasonalization by fitting a trend and then averaging the seasonal factors using *medians* instead of *means,* so the results are not exactly the same as with MegaStat. Minitab offers excellent graphical displays for decomposition, as well as forecasts, as shown in Figure 14.30. Minitab offers additive as well as multiplicative seasonality. In an additive model, the *CMA* is calculated in the same way, but the raw seasonals are *differences* (instead of ratios) and the seasonal indexes are forced to sum to *zero* (e.g., months with higher sales must exactly balance months with lower sales). Most analysts prefer multiplicative models for trended data.

Figure 14.30

Minitab's Graphs for Floor Covering Sales **FloorSales**



Fitting Trends in R **FireLosses**

We can fit and display several trend models (linear, quadratic, exponential) on the same graph using the 8 years of fire loss claims (Table 14.10). First, create a vector Claims containing the data and a vector Time as an index for the years 1, 2, . . . , 8. Put the variables in a data frame CData. Print the data to verify it:

```
> Claims = c(12.940, 11.510, 12.428, 13.457, 17.118, 17.586, 21.129, 18.874)
> Time=c(1:8)
> CData=data.frame(Claims,Time) # create a data frame for convenience
> CData                          # list the data
```

	Claims	Time
1	12.940	1
2	11.510	2
3	12.428	3
4	13.457	4
5	17.118	5
6	17.586	6
7	21.129	7
8	18.874	8

Fit the trend models. Save each result for later use. Display their fitted predictions so you can compare them. In this case, predictions are similar, which favors the simpler linear model:

```
> fit1=lm(Claims~Time)             # fit linear trend (1st deg poly)
> fit2=lm(Claims~poly(Time,2))     # fit quadratic trend (2nd deg poly)
> fit3=lm(log(Claims)~Time)        # fit exponential trend in logs

> fitted.values(fit1)
        1        2        3        4        5        6        7        8
 11.09825 12.39311 13.68796 14.98282 16.27768 17.57254 18.86739 20.16225

> fitted.values(fit2)
        1        2        3        4        5        6        7        8
 11.61575 12.46704 13.46618 14.61318 15.90804 17.35075 18.94132 20.67975

> fitted.values(fit3))
        1        2        3        4        5        6        7        8
 11.44177 12.43040 13.50446 14.67132 15.93901 17.31623 18.81245 20.43795
```

Save the predictions for each model. Plot the data and all 3 fitted curves on the same graph in colors with variable names as axis labels. Use exp() for exponential since its predictions are logs. Separately, plot the preferred linear model, using its fitted equation as a title (rounded to 3 or 4 decimals for brevity):

```
> Pred1=predict(fit1,data=CData)           # linear trend
> Pred2=predict(fit2,data=CData)           # quadratic trend
> Pred3=exp(predict(fit3,data=CData))      # exponential trend
> plot(Time,Claims,main=
  "Three Trend Models"                     # plot x,y data
> lines(Time,Pred1,col="red")              # plot linear trend
> lines(Time,Pred2,col="blue")             # plot quadratic trend
> lines(Time,Pred3,col="green")            # plot exponential trend
> b0=round(fit1$coefficients[[1]],3)       # plot linear separately
> b1=round(fit1$coefficients[[2]],3)
> r2=round(summary(fit1)$r.squared,4)
> B1=as.character(abs(b1))
> B0=as.character(b0)
> R2=as.character(r2)
> if(b1<0) DispEqn=paste("y =",B0,"-",B1,"x ","Rsq =",R2) else
+  DispEqn=paste("y =",B0,"+",B1,"x ","Rsq =",R2)
> plot(Time,Claims,main=DispEqn)
> lines(Time,Pred1,col="red")
```

If you need more details (R^2, coefficients, standard error, t-values, p-values, residuals) you can use the summary() function for any of the fitted models. For example, here is a summary for the linear model:

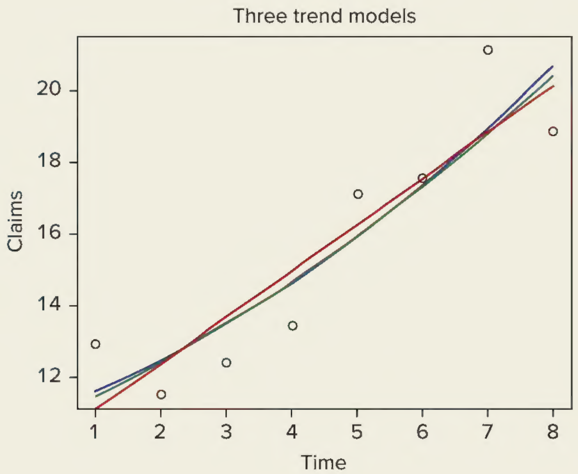
```
> summary(fit1)
Residuals:
    Min       1Q   Median      3Q     Max
-1.5258  -1.2670  -0.4348  1.0907  2.2616

Coefficients:
            Estimate Std.  Error t value Pr(>|t|)
(Intercept)  9.8034   1.2551   7.811   0.000232 ***
Time         1.2949   0.2486   5.210   0.001996 **
---
Signif. codes:  0 '***'  0.001 '**'  0.01 '*'  0.05 '.'  0.1 ' '  1
Residual standard error: 1.611 on 6 degrees of freedom
Multiple R-squared: 0.8189, Adjusted R-squared: 0.7888
F-statistic: 27.14 on 1 and 6 DF, p-value: 0.001996
```

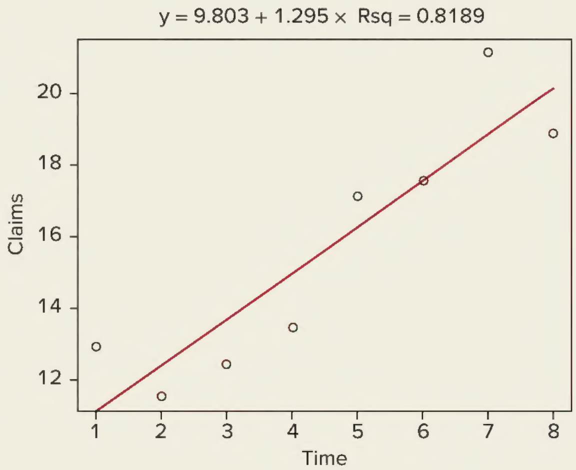
Seasonal Decomposition in R FloorSales

It is easy to calculate seasonal factors for monthly or quarterly time series data. For example, we have an Excel spreadsheet with quarterly floor sales data (24 quarters starting in year 2011) as shown in textbook Table 14.14:

Year	Quarter	Sales
2012	Qtr 1	259
	Qtr 2	236
	Qtr 3	164
	Qtr 4	222
	…	…
	…	…
	…	…
2017	Qtr 1	626
	Qtr 2	535
	Qtr 3	397
	Qtr 4	488

Import the Sales column (including the heading) into a data frame called FloorSales (we ignore the year and quarter labels):





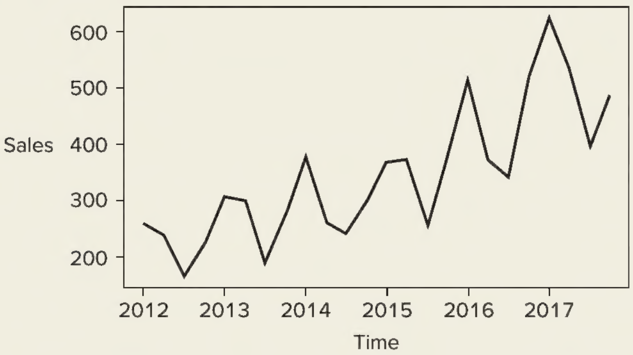
```
> FloorSales=read.table(file="clipboard",sep="t",header=TRUE)
> FloorSales
        Sales
1       259
2       236
3       164
4       222
.        .
.        .
.        .
21      626
22      535
23      397
24      488
```

Use the ts() function to convert the data to a quarterly time series with frequency=4 and year labels beginning with 2012 using start=2011. Plot the data.

```
> ts.FloorSales=ts(FloorSales,frequency=4,start=2012)
> ts.FloorSales
```

	Qtr1	Qtr2	Qtr3	Qtr4
2012	259	236	164	222
2013	306	300	189	275
2014	379	262	242	296
2015	369	373	255	374
2016	515	373	339	519
2017	626	535	397	488

```
> plot(ts.FloorSales)
```



Use the decompose() function to obtain seasonal factors. The results match Table 14.15 in the textbook and show all the steps of decomposition. R also displays the deseasonalized time series and residuals around the fitted trend. The last line shows the four seasonal factors (they are adjusted so they sum to 1), which are identical to those in Table 14.16:

```
> decompose(ts.FloorSales,"multiplicative")
```

$x

	Qtr1	Qtr2	Qtr3	Qtr4
2012	259	236	164	222
2013	306	300	189	275
2014	379	262	242	296
2015	369	373	255	374
2016	515	373	339	519
2017	626	535	397	488

$seasonal

	Qtr1	Qtr2	Qtr3	Qtr4
2012	1.2520398	1.0210404	0.7396682	0.9872516
2013	1.2520398	1.0210404	0.7396682	0.9872516
2014	1.2520398	1.0210404	0.7396682	0.9872516
2015	1.2520398	1.0210404	0.7396682	0.9872516
2016	1.2520398	1.0210404	0.7396682	0.9872516
2017	1.2520398	1.0210404	0.7396682	0.9872516

$trend

	Qtr1	Qtr2	Qtr3	Qtr4
2012	NA	NA	226.125	240.000
2013	251.125	260.875	276.625	281.000
2014	282.875	292.125	293.500	306.125
2015	321.625	333.000	361.000	379.250
2016	389.750	418.375	450.375	484.500
2017	512.000	515.375	NA	NA

$random

	Qtr1	Qtr2	Qtr3	Qtr4
2012	NA	NA	0.9805242	0.9369445
2013	0.9732252	1.1262787	0.9237053	0.9912849
2014	1.0701053	0.8783946	1.1147316	0.9794112
2015	0.9163438	1.0970380	0.9549839	0.9988911
2016	1.0553657	0.8731728	1.0176267	1.0850399
2017	0.9765315	1.0166876	NA	NA

$figure

```
[1]  1.2520398    1.0210404    0.7396682    0.9872516
```